OM of: US-09-652-292-2 to: Issued_Patents_NA:*   out_format : pfs

Date: Mar 15, 2002   7:54 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09652292/runat_13032002_161725_3228/app_query.fasta_1.606
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPCL=0.000
-QGAPEXT=4.500 -QGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09652292_@CGN1_1.71 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-652-292-2
Query length: 541
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 113238999
Search time (sec): 48.280000

/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-335-409-1	+	134.50	147.92	1.37	68750
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-08-964-127-1	+	133.50	187.88	0.0081	2460
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-038-832-1	+	125.50	171.48	0.0667	2900
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-400-159-7	+	124.50	165.25	0.1485	4464
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-611-729A-7	-	124.50	165.19	0.1495	4483

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-103-840A-1

seq_documentation_block:
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

alignment_scores:
           Quality: 457.50   Length: 528
           Ratio: 1.628   Gaps: 15
  Percent Similarity: 53.220   Percent Identity: 27.652

alignment_block:
US-09-103-840A-1 x US-09-103-840A-1   .
Align seg 1/1   to: US-09-103-840A-1   from: 1   to: 4411529

    15 SerLeuLeuGlyGlyLeuThr.....................PheGlyTy 24
       :::|||||| |||||||| :::||||||
3717235 GCGACYCCTCGTCGGGCTCACCGGCGCCAGCGTCGGCGTCCTCTACCGGTA 3717285

    24 rGluLeuAlaValIleSerGlyAlaLeuLeuProLeuGlnLeuAspPheG 41
       |:::||||:::||||:::||||||||||||||| |||::| :::||||
3717286 CGAACCTTTCCGCCATCGGCGGGTGCCTGGCTGCTGTCTCCAGCCAGGAATTCG 3717335

    41 lyLeuSerCysLeuGluGlnGluPheLeuValGlySerLeuLeuGly 57
       |||: ||||||||||||||:::||||| :::||||||
3717336 AACTCACCACTCGAGAACAGGAGTTGCTAGCCACCACGGCGTGCTCGGC 3717385

    58 AlaLeuLeuAlaSerLeuValAlGlyPheLeuIleAspCysTyrGlyAr 74
       ::: ||||||| ||||:::   ::||| ||||
3717386 CAGATCGCCGGGGCGGCCTTGGGCGGCCATCCTCGCCAACGCGATCGGACG 3717435

    74 gLysGlnAlaIleLeuGlySerAsnLeuValLeuLeuAlaGlySerLeuT 91
       |||| :::|||:::|||||||||||
3717436 CAAGAAATCGGTG.....................GTGCTCATCGTCGCCGGCTACGCAG 3717473

    91 hrLeuGlyLeuAlaGly..........SerLeuAlaTrpLeuValLeu 103
       :::|||:::||| ||| ||||:::|||||||
3717474 TGTTCGCCCCTGCTCGCCGGCGACCTCGGTCCGTACCGATCCTGGTGGTG 3717523

   104 GlyArgAlaValAlGlyPheAlaIleSerSerLeuSerMetAlaCysCy 120
       :::||| ||||:::|||| ::|||||||||:::||||||||||
3717524 GCGCGTCGTCGGTGTGACAATCGGCCTGTCGGTGGTGGTGGTGCC 3717573

   120 siIeTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuVals 137
       :::|||:::|||:: |||::| |||:||:||
3717574 AGGTGTATGTGGCCGAGTCGGCGCCGGGCCGGGCTGGGTCGTTGGTGA 3717623

score_list:
Sequence	Strd	Orig ZScore	EScore Len	Documentation			
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-103-840A-1	+	457.50 640.73		4.9e-28	4411529	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-591-025-8	+	381.50 605.57		4.4e-26	2592	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-928-692-9	+	314.50 489.88		1.2e-19	3000	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-09-031-392-1	+	284.50 443.21		4.9e-17	2343	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-09-299-549-1	+	284.50 443.21		4.9e-17	2343	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-501-572-6	+	226.50 348.02		9.8e-12	1896	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-040-444-6	-	226.50 348.02		9.8e-12	1896	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-501-572-4	+	201.00 305.09		2.4e-09	1882	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-040-444-4	-	201.00 305.09		2.4e-09	1882	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-501-572-5	+	198.50 300.86		4.2e-09	1885	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-040-444-5	-	198.50 300.86		4.2e-09	1885	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-103-840A-2	-	195.00 197.95		0.0022	4403765	1	
/cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-676-967-5	-	167.50 246.20		4.6e-06	2277	1	
/cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-676-974-5	-	167.50 246.20		4.6e-06	2277	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-09-098-487-5	-	167.50 246.20		4.6e-06	2277	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-743-637B-182	+	163.00 245.29	5.2e-06	1335			
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-647-397-1	+	160.00 234.55		2.0e-05	2102	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-118-442-23	+	158.50 255.06		1.5e-06	333	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-677-064-23	+	158.50 255.06		1.5e-06	333	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-403-852P-6	+	155.50 228.67		4.4e-05	1833	1	
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-510-646B-6	+	155.50 228.67		4.4e-05	1833	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-231-818-6	+	155.50 228.67		4.4e-05	1833	1	
/cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-125-468-1	-	145.50 176.85		0.0336	30001	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-474-933-1	-	145.50 176.85		0.0336	30001	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-08-994-035C-4	-	144.50 193.01		0.0042	7198	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-103-840A-1	-	142.00 108.53		214.30	4411529	1	
/cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US96-10986-13	-	140.50 207.97	0.0006	1269			
/cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-882-046-3	-	140.50 192.66		0.0044	4315	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-494-907-2	-	140.50 190.63		0.0057	5076	1	
/cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US96-10986-2	-	140.50 190.63	0.0057	5076			
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-494-907-4	-	140.50 188.19		0.0078	6170	1	
/cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US96-10986-4	-	140.50 188.19	0.0078	6170			
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-494-907-3	-	140.50 187.76		0.0083	6387	1	
/cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US96-10986-3	-	140.50 187.76	0.0083	6387			
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-752-760A-1	-	139.50 164.77		0.1579	35081	1	
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-998-416-719	-	138.00 211.05		0.0004	708	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-020-956-110	+	136.50 188.86		0.0072	3410	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-030-607-110	-	136.50 188.86		0.0072	3410	1	
/cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-443-501A-2	+	136.50 150.72		0.9573	71989	1	

137 erLeuTyrGluAlaGlyIleThrvalGlyIleLeuLeuSerTyrAlaLeu 153
      :: :: ||||||:: :: ||||||||||||||||||
3717624CCGGCTATCAGCTGGCGACGCTTAGCCGCATCGTCGTCGGTTACCTGGTC 3717673

154 AsnTyrAlaLeuAlaGlyThrProGlyTrpArgHisMetPheGlyTr 170
      |||||| ||||||||::: ||||||||
3717674GGCTACCTGTTGGCCGGATCG...CACGGCGTGGCGCCGATGTTCGGGCT 3717720

170 pAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheLeuProA 187
      ||||||||||||||||||::: ::   ::
3717721GGCCACGGCCGCGCCGCCACGCTG...CTGTTGCCGTTGTTGTGGCG..... 3717762

187 laGlyThrAspGluThrAlaThrHisLysAspLeuProLeuGlnGly 203
      ||| :::::: ||||: |||| :::
3717763.......CATGCCGATACCGCCCGCTGGTATCTGCTCAAGGG 3717798

204 GlyGlu........................AlaProLysLeuGlyProGlyAr 213
      ::: ||| |||||
3717799CCGGATGCCGACGCGCGCTAGCGCGTCGCGGGATCCAGCCGGAGGCCG 3717848

213 gProArg.TyrSerPheLeuAspLeuPheArgAlaArgAspAsnMetArg 229
      |||:::: || :::: |||
3717849ACATCGATGCCGAGCTGGCCGATATGCGGCCGCGGTCGCAACGGCGGC 3717898

230 Gly............................ArgThrThr...... 233
      |||
3717899GGCGGTATCGGCGGAAATGGTCGGCGGCCGTATCTCGCGGCCACGCTGTT 3717948

234 ...ValGlyLeuGlyLeuValLeuPheGlnLeuThrGlyGlnPro A 249
      ::::::::||||||:::::
3717949CGTCATCGGCGTCGGCTTCCCTGTC......CAGATCACCGGGATCAACG 3717992

249 snValLeuCysTyrAlaSerThrIlePheSerValGlyPheHisGly 265
      :::::: :: ||||||||| ::: |||||| |||
3717993CGATCATCTACTACAGTCCGCACTTTTCGCCGCCATGGGCTTCGCGGGC 3718042

266 GlySerSerAlaValAlaValGlyLeuGlyAlaValLysValAl 282
      ::: :: :: |||||||||||||
3718043TATTTCGGCGATGCTTGCCCTGCCCGCATGGTGCAAGTCGCCGGCTTGC 3718092

282 aAlaThrLeuThrAlaMetGlyLeuValAspArgAlaGlyArgArgAlaL 299
      |||| ::::::: |||||||||| |||||
3718093GGCGCGGTGTGTCCGCTCGCTGTTTCTGTCGATCGGCTGGCCGTCGCCCGA 3718142

299 euLeuLeuAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIleGly 315
      :: |||||||||::: ||| ||| :::::::::
3718143TCCTGTTGTGCCGGACGATG...ATCACCCGAGATGCCGTGCTG 3718189

316 LeuValSerPheAlaValProMetAspSerGlyProSerCysLeuAlaVa 332
      ::::: ||||||||||||||
3718190ATCACCGTATTCGCCAACGACTCCGATGGTGGC...................... 3718222

332 lProAsnAlaThrGlyGlnThrGlyLeuPro...GlyAspSerGlyLeuL 348
      ||||||||| ||| :::||:::
3718223...............ACGGGCTGGTGTTGGGGTTCGCCGGCCGTGC 3718253

348 euGlnAspSerSerLeuProProIleProArgThrAsnGluAspGlnArg 364
      ||
3718254TG.. 3718255

365 GluProIleLeuSerThrAlaLysLysThrLysProHisProArgSerGl 381
3718255.. 3718255

381 yAspProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGlyP 398
3718255.. 3718255

398 roProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeu 414
3718255.. 3718255

415 CysLeuMetValPheValSerAlaPheSerGlyPheGlyProValTh 431
3718255.. 3718255

... CTGTTCATCATCGGTCGGTTCAACTTCGATTCGGCTCGCTGCTGGT 3718296

431 rTrpLeuValLeuSerGluIleTyrProValGluIleArgGlyArgAlaP 448
      ::||| |||| ::||::: ::||::: :::
3718297CTGGGTGCTACGCCCGGAGAGCTTCCCGTCCGGCTGCGTGCGGTCGATGGGAT 3718346

448 heAlaPheCysAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeu 464
      ::: ||||::::::
3718347CGAGCCCGATGCTCACCTCGACACTGACGGGCCACGCGACCGTTGCCGCC 3718396

465 SerPheLeuAspLeuPheGlyThrIleGlyLeuSerTrpThrPheLeuLe 481
      ||| ::::::: |||| ::: ::: :::
3718397TTCTGCGTCACGATGCTGCCTGTGTCCGGCGGCGCCAGGCGTTTTCGCGGT 3718446

481 uTyrGlyLeuThrAlaValLeuGlyLeuGlyIleLeuTyrLeuPheValP 498
      ::: ||| ||||||:::::   ::||| |||::::
3718447CTTCGGCACGTTCGCCGTGCGGTTCGTGCTGTACCGCTTTGGC 3718496

498 roGluThrLysGlyGlnSerLeuAlaGluIle 508
      |||||||||||| |||||
3718497CGGAGACCAAGGGCCGCAAACTCGAGGAGATC 3718528

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-591-025-8

alignment_scores:
         Quality: 381.50   Length: 486
         Ratio: 1.467   Gaps: 13
Percent Similarity: 53.498   Percent Identity: 26.543

alignment_block:
US-09-652-292-2 x US-09-591-025-8

Align seg 1/1 to: US-09-591-025-8 from: 1 to: 2592

56 LeuGlyAlaLeuAlaSerLeuValGlyValGlyPheLeuIleAspCysTy 72
     ::||||::||::::::||||::: |||:::: ||||:::
583 GTGGGCGGCATGATTTCCTCCTTCCTCATTGGTATCATCTCTCAGTGGCT 632

72 rGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuAlaGlyLys 89
     |||||||||||| ::||:::||| ::: ||||:::||
633 TGGAAGGAAAAGGCCATTGCTGGTCAACAATGTCCTGGCGGTGCTGGGGG 682

89 erLeuThrLeuGlyLeuAlaGlySerLeuAla........TrpLeuVal 102
     :: :||||||||||::: |||
683 GCAGCCTCATGGGCCTGGCCAACCGTCGTGCCTCCTATGAAATGCTCATC 732

```
 103 LeuGlyArgAlaValValGlyPheAlaIleSerLeuSerMetAlaCy                            119
 733 CTTGGACGATTCCTCATTGGCGCTCCTACCTCAGGGCTGACATCAGGGCTGGT                      782

 119 sCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuV                          136
 783 GCCCATGTACCTGGGGGAGATTGCTCCCACTCACCTGCGGGCGCCCTGG                          832

 136 alSerLeuTyrGluAlaGlyIleLeuThrValGlyIleIleLeuLeuSerTyrAla                    152
 833 GGACGCTCAACCAACTGGCCATTGTTATCCGGCATTCTGATCGCCCAGGTG                        882

 153 LeuAsnTyr...AlaLeuAlaGlyThrProTrpGlyTrpArgHisMetPh                          168
 883 CTGGGCTTGGAGTCCCTCCTGGGACTGCCAGCCTGTGGCCACTGCTCCT                          932

 168 eGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuPheL                            185
 933 GGGCCCTCACAGTGCTACCTGCCCTCCTGCAGCTGGTCCTGCCCTTCT                          982

 185 euProAlaGlyThrAsp..............GluThrAla                                   193
 983 GTCCCCGAGAGCCCCGCTACCTCACCTCTACATCATCCAGAATCTCGAGGGGCCT                   1032

 194 ThrHisLysAspLeuIleProLeuGlnGly..............                               203
1033 GCCAGAAAGAGTCTGAACGCCTGACAGCCTGGGCCGATGTTTCTGGAGT                         1082

 204 .........GlyGluAlaProLysLeuGlyProGlyProArgProA                             215
1083 GCTGGCTGAGCTGAAGGATGAGAAGCGGAAGCTGGAGCGTGAGCGGCCA.                        1131

 215 rgTyrSerPheLeuAspLeuPheArgAlaArgAspAsnMetArgGlyArg                          231
1132 ..CTGTCCCTGCTCCAGCTCCTGGGCAGCCGTACCCAC...CGGCAGCCC                        1176

 232 ThrThrValGlyLeuGlyLeuValLeuPheGlnLeuThrGlyGlnPr                            248
1177 CTGATCATTGCGGTCGTGCTGCAGCTGCTCTCTGGCATCAA                                 1226

 248 oAsnValLeuCysTyrAlaSerThrIlePheSerSerValGlyPheHisG                          265
1227 TGCTGTTTTCTATTATTCGACCAGCATCTTCCAGACAGCAGGGGTA....                        1272

 265 lyGlySerSerAlaLeuAlaSerValGlyLeuValLeuValLysVal                            281
1273 .....GGCCAGCCTGCCTATGCCACCATAGGAGCTGGTGTGGTCAACACA                        1317

 282 AlaAlaThrLeuThrAlaSerMetGlyLeuValAspAlaArgAlaGlyArgArgAl                    298
1318 GTCTTCACCTTGGTCTCGGTGTTCTTGGTGGAGCGGCGGGGGCGCCGGAC                        1367

 298 aLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIleG                          315
1368 GCTCCCATCTCCTGGGCCTGGCG.............G                                      1390

 315 lyLeuValSerPheAlaValProMetAspSerProSerCysLeuAla                            331
1391 GCATGTGTGGCTGTGCCATCCTGATGACTGTGGCTCTGCCTCCTGCTG...                       1437

 332 ValProAsnAlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuLe                          348
1437 ......................................                                    1437

 348 uGlnAspSerSerLeuProProIleProArgThrAsnGluAspGlnArgG                          365
1438 ..............................GAGCCGAG                                    1444

 365 luProIleLeuSerThrAlaLysLysThrLysProHisProArgSerGly                          381
1445 TTCCCCATGAGC......................                                         1458

 382 AspProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGlyPr                          398
1458 ....................                                                      1458

 398 oProLeuProAlaArgGlyHisAlaLeuLeuLeuLeuArgTrpThrAlaLeuLeuC                    415
1459 ..................TACGTCTCCATTGTGG                                        1474

 415 ysLeuMetValPheValSerAlaPheSerPheGlyProValThr                               431
1475 CCATCCTTGGCTTCGTGGCATTTTTGAGATTGGCCTGGCCCCCATTCCT                         1524

 432 TrpLeuValLeuSerGlyIleTyrProValGluIleArgGlyAlaIleAlaPh                       448
1525 TGGTTCATCGTGCCCAGCTCTTCAGCCAGGGACCCCGCCCCCGCAGCCAT                        1574

 448 eAlaPheCysAsnSerPheAsnTrpAlaAlaAlaAsnLeuPheIleSerLeuS                       465
1575 GGCTGTGGCTGGTTTCTCCAACTGGACAGCAACTTCATCATTGGCATGG                         1624

 465 erPheLeuAspLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeu                          481
1625 GTTTCCAGTATGTTGCGGAGGCTATGGGG...CCCTACGTCTTCCTTCTA                        1671

 482 TyrGlyLeuThrAlaValLeuGlyLeuGlyPheIle.........TyrLe                          495
1672 TTT.......GCGGTCCTCCTGCTGGGCTTCTTCATCTTCACCTTCTT                          1712

 495 uPheValProGluThrLysGlyGlnSerLeuAlaGluIleAspGlnGlnInP                        512
1713 AAGAGTACCTGAAACTCGAGCCCGGACGTTTGACCAGATCTCGGCTGGCCT                       1762

 512 heGlnLys                                                                   514
1763 TCCACCGG                                                                  1770
```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-928-692-9

seq_documentation_block:
; Sequence 9, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
;   APPLICANT: Brody, Howard
;   APPLICANT: Yaver, Deborah S.
;   APPLICANT: Lamsa, Michael
;   APPLICANT: Hansen, Kim
;   TITLE OF INVENTION: Methods for Modifying the Production of
;   TITLE OF INVENTION: a Polypeptide
;   NUMBER OF SEQUENCES: 80
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
;   STREET: 405 Lexington Avenue
;   CITY: New York
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 10174
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/928,692
;   FILING DATE: 12-SEPT-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Lambiris, Elias J
;   REGISTRATION NUMBER: 33,728
;   REFERENCE/DOCKET NUMBER: 4944.200-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 212-867-0123
;   TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:

```
;    LENGTH:  3000 base pairs
;    TYPE:  nucleic acid
;    STRANDEDNESS:  single
;    TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-928-692-9

alignment_scores:
        Quality:     314.00          Length:      591
        Ratio:       1.150           Gaps:        18
Percent Similarity:  46.193   Percent Identity:  23.181

alignment_block:
US-09-652-292-2 x US-08-928-692-9    .
Align seg 1/1  to: US-08-928-692-9   from:  1  to: 3000
```

```
  2 GlyHisSerProProValLeuPro.LeuCysAlaSerValSerLeuLeuG   18
    ||||||||| |||  ||||||| ::          :::::  ||||||
557 GGTCACTCCGTACCTGACCTACCTGTG......TTTATCACAACTTTGG   600

 18 lyGlyLeuThrPheGlyTyrGluLeuValIleSerGlyAlaLeuLeu     34
    || ||| ||||||||:::|||       ||||:  ||||
601 GGCCACTTCAATTCGGATATCATTTG...GTATTACACGGAGCTTGGTCT   647

 35 ProLeuGlnLeuAspPheGlyLeuSer..................       43
    ||:::   |||       |||
648 ATGCTGGAGGC...TTCAATACATCGGCTGACAATATATTATGATAGGC   693

 43 ..............................CysLeu.........Glu    46
                                  |||:::          :::
694 TGAGCTCAATGCCCCCCAGCCGTGATAACTTGCGAGCGGAAAAGCATCC   743

 44 ...................................CysLeu........Glu   46
                                       |||:::
744 ATTCGACAACAACACGGGTCTCCCGCAATGCATACCTATGAACCCATCC   793

 47 GlnGluPheLeuValGlySerLeu...LeuLeuGlyAlaLeuLeuAlaSe   62
    |||   ||||||:::||||: :        ||||||:::|||||:||::
794 CAATTCGGCCTGGTCTCCTCTATATACACCCTGGGGGCTTGCTAGGGGC   843

 62 rLeuValGlyGlyPheLeuIleAspCysTyrGlyArgLysGlnAlaIleL   79
    :||||:||::: :::        |||      ||| ::::  ||||||
844 TCTCCTGGCAGGTCCAGTTTCCACCAAGCATGGCCCGTTGTTCACACTGC   893

 79 euGlySerAsnLeuValLeuLeuAlaGlySerLeuThrLeuGlyLeuAla   95
    :::::::::::  ::::::  |||         :::  :::  ||||
894 GAGCGACCACCATCTTCTTCATCCTAGGCCCTATAGCAGAAAACATTTGCG   943

 96 GlySerLeuAlaTrpLeuValLeuGlyArgAlaValValGlyPheAlaIl   112
    ||:::     |||    :::|||||| ::  |||  :::
944 CCCGATATACCCCGTATTGAGTATGGGTAGGCTTTTATCTGGTGTTGGTGC   993

112 eSerLeuSerSerMetAlaCysCysIleTyrValSerGluLeuValGlyP   129
    ::  |||     ::||||||  ||::|||::||||||||||:::|||||  |
994 GGGCGCTTCTATCGTGTGGGGTCCGATATATATCTCTGAGATTGCTCCTC   1043

129 roArgGlnArgGlyValLeuValSerLeuTyrGluAlaGlyIleThrVal   145
    :::||:::|:::::::  ::|||::           :::|||
1044 CTAGTGCTAAGGGTCTTTTCGGCGCTTTTACGCAAATCATGACTAATGTC   1093

146 GlyIleLeuSerTyrAlaLeuAsnTyrAlaLeuLeuAlaGlyThrProTr   162
    |||||||||||:::|||||  :::||||||  |||||  :::|||||||
1094 GGTATTCGTTGACACAGTCCCTTGGTTACTTCTTGAGTAAAGGAAGTAT   1143

162 pGlyTrpArgHisMetPheGlyTrpAlaThrArgProAlaValLeuGlnS   179
    :::||:||: |||      ||| ||| |||:::||
1144 G...TGGAGAGTTATACTTGCAATTGCTGGCGCGGATCCGATGCCTTGAGC   1190

179 erLeuSerLeuPheLeuProAlaGlyThrAspGluThrAlaThrHis     195
    ||||||:::::||:::||||||    ||  ||| ||||||
1191 TTCTGGGCCTCTTCTTTAGTCCCAGAAAGCCCCATCTGCCTTGCAGATCAC   1240
```

```
196 LysAsp.............................................  197
    :::
1241 CAGAAAGGAATGTGGCTAGACAGGTGCTACAACGTATACGGGGCAGGGA  1290

197 ................................................... 197

1291 TGCAGACATCGAGCCACCAGAGGTTGAAGGCTGGAGAACATCTGCAGCGCCTG  1340

198 .................LeuIleProLeuGlnGlyGly              204
                     |||   |||  :::|||:::
1341 AACACAGCTCTGGGAGAGCAGTCCCTACTATCACCCCATCTGGAAAT    1390

205 GluAlaProLysLeuGlyLeuProGlyArgProArgTyrSerPheLeuAspLe  221
    ||||||     ::::::  ||||||             |||
1391 ATGCCACCCAAGCAACCTCCGGTTACCATGATGCGAGCTATTACTGATTC  1440

221 uPheArgArgAspAsnMetArgGlyArgThrThrValGlyLeuLeuGlyL  238
    |||             |||                  :::::  |||
1441 TTTT...........TACCCGCCTGCCATCATTGCAGTGGTCGGAG    1475

238 euValLeuPheGlnGlnLeuThrGlyGlnProAsnValLeuCysTyrAla  254
    ::::::::  ||||||||||:::|||||     :::::::    |||
1476 TCATGGTTCCCAGCAGTTCACTGTGTCAACAGCATCATCATGTACAGC   1525

255 SerThrIlePheSerValGlyPheHisGlyGlySerAlaValLe        271
    :::::::::||        :::::::::: |||
1526 GTTTCCCTCTTACAGACCATC.........CTTCCCACCACTGCAGCCCT  1566

271 uAlaSerValGlyLeuGlyAlaValLysValAlaAlaAlaThrLeuThrAlaM  288
    | |||||   ||||||||::: :::|||||:::::::::
1567 GTTGTCGGTGATCATCTCGGCTATCAATCTTGTAATCACTCTGGCCTGCT   1616

288 etGlyLeuValAspArgArgAlaGlyArgArgAlaLeuLeuLeuAlaGlyCys  304
    |||          ||||||||||:::  |||||
1617 CACCACTACCTGATAAGATTGGTAGACGCTCCTGCCTGCTT........   1657

305 AlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValSerPheAlaVa  321
    ::::::::::  :::|||:::||||||::|||||  |||
1658 .........CTAAGTATCAGCGCATGGGTCTTAATTCCGTCCTACT    1695

321 lProMetAspSerGlyProSerCysLeuAlaValProAsnAlaThrGlyG  338
    :::::::||||||  :::||||||||
1696 G.............................GCGCTAGCCATC.........  1708

338 lnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeuPro  354
    :::||||||::::::::::::::::  :::::
1708 ...................................................  1708

355 ProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerThrAl  371
    :::::
1708 ...................................................  1708

371 aLysLysThrLysProHisProArgSerGlyAspProSerAlaProProA  388
    :::::                 |
1708 ...................................................  1708

388 rgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaArgGly  404
    :::::
1708 ...................................................  1708

405 HisAlaLeuArgTrpThrAlaLeuLeuCysLeuMetValPheValSe    421
    :::::          ::  ::::::::::::::::::|||||:
1709 TACTTCAACCTGAAAGCCTTATCCGCCATAGCAGTTCTACTTTCGTTGC  1758

421 rAlaPheSerPheGlyPheGlyProValThrTrpLeuValLeuSerGluI  438
    ::::||:  |||:::::::::|||:::  |||||:::||:||:
1759 TTCTTTCCCGCGCCGTCTAGGCCCAGTCCCCTCATTTAGCCTCTGAAC   1808

438 leTyrProValGluIleArgGlyAlaPheAlaPheCysAsnSerPhe     454
    ::  |||    |||       |||  ||
1809 TCGTTGCCCCGGAGGCTGTCGGCGGCCGACAGAGCTGGGCGCTGGGAGCG   1858
```

```
  10 LeuCysAlaSerValSerLeuLeuGlyGlyLeuThrPhe.......GlyTy  24
     |||  ||::|||  |||||||:::::::  ||: |||  ::|||      |||||
 241 CTCGTGGCCTCCGTCCTCGGCGGCCTTCGGCTCCTTCCTCCTACGGCTA  290

  24 rGluLeuAlaValIleSerGly............................  31
     |::||:|:::|||  |||  :::                           :::
 291 CAACCTTCGGTGGTGAATGCCCCACCCCGTACATCAAGGCCTTTTACA  340

  32 ..................................AlaLeuProLeuGlnLeu  38
                                        ||||||||||||||
 341 ATGAGTCATGGGAAAGAAGGCATGGACGTTCCAATAGACCCAGACACTCTG  390

  39 AspPheGlyLeuSerCysLeuGluGlnLeuGlnPheLeuGlyLeuSerLeuLe  55
     :::       |||      |||||   ||  ::| |||  :::||::: ||
 391 ACTCTTGCTCTGCTCTGTGACTGTCTCCATATTCGCCATCGGTGGA...CT  437

  55 uLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleAspCysT  72
     |:::||:||||:|||||::::||||||  |||       |||
 438 TGTGGGGACGTTAATTGTGAAGATGATTGGAAAGGTTCTT..........  477

  72 yrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeuAlaGly  88
     |||||||||  ||:||||:|||  |||::::|||
 478 ..GGGGAAGCACACTTGCTGCCAATAATGGGTTGCAATTTCTGCT  525

  89 SerLeuThrLeuGlyLeu.........AlaGlySerLeuAlaTrpLeuVa  102
     ::::||  ||||||||||             |||||||:::
 526 GCATTGCTGATGGCCTGCCTCCGCTCCAGGCAGGAGCCTTTGAAATGCTCAT  575

 102 lLeuGlyArgAlaValAlaGlyPheAlaIleSerSerLeuSerMetAlaC  119
     :::|||||  :::||:||||:::::|           :::
 576 TGTGGGACGCTTCATCATGGGCAATAGATGGAGGCGTTCGCCCTCAGTGTGC  625

 119 ysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeu  135
     :::||||||:::||||||::    |||         :::||||    ||:
 626 TCCCCATGTACCTCAGTGAGATCTCACCCAAGGAGATCCGTGGCTCTCTG  675

 136 ValSerLeuTyrGluAlaGlyIleLeuThrValGlyIleLeuLeuSerTyrAl  152
     |||:::              |||    ::||||||:|||:::
 676 GGGCAGTGACTGCCATCTTTATCTGCATTGGCGTGTTCACTGGGCAGCT  725

 152 aLeuAsnTyr...AlaLeuAlaGlyThrProTrpProGlyTrpArgHisMetP  168
     ||:::             ||:||::|||||||    |||  |||
 726 TCTGGGCCTGCCCGAGCTGCTGGGAAAGGAGAGTACCTGGCCATACCTGT  775

 168 heGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPhe  184
     ||||||||:::::|||||||||||    |||:::
 776 TTGGAGTGATTGTGGTCCCTGCCGGTTGTCCAGCTGCTGAGCCTTCCCTTT  825

 185 LeuProAlaGlyThrAspGluThrAlaThrHisLysAspLeuIleProLe  201
     |||
 826 CTC...........................................  828

 201 uGlnGlyGlyGluGluAlaProLysGlyProGlyLeuGlyProArgTyr....  216
     |||  |||||||||||
 829 .................CCGGACAGCCCAGCGTACCTGC  850

 216 .................................................  216
 851 TCTTGGAAGCACAACCAGGCAAGAGCTGTGAAAGCCTTCCAAACGTTC  900

 216 .................................................  216
 901 TTGGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAGGTCCTGGCTGAGAG  950

 217 .........................................SerPheLeuAspLeuPheArgA  224
                                              |||  ||::||| ||:::||||
 951 CCACGTGCAGAGGAGCATCCGCTGGTCGTCGTGGAGCTGCTGAGAG  1000

 224 laArgAspAsnMetArgGlyArgThrValGlyLeuValLeu  240
     ||  |||  :::              :::    ::|
1001 CT...CCCTACGTCCGCTGGCAGGTGGTCACCGTGATTGTCACCATGGCC  1047

 241 PheGlnIleLeuThrGlyGlnProAsnValLeuCysTyrAlaSerThrIl  257
```

```
 455 AsnTrpProAlaAlaAsnLeuPheIleSerLeuSerPhe........LeuAs  468
     ||||||            |||::::::||             |||           ||
1859 AACTGGATTGGCACGTTCATCGTGGCACAATTTTTTCCGATGTTAAACGA  1908

 468 pLeuIleGlyThrIleGly...LeuSerTrpPheThrPheLeuTyrGlyL  484
     |||||||||::||||||         |||    ||  |||  |||   ::::
1909 TTTGTTGGGCGGACGAGGCAAGATCTACTGGATCTTTGCA......GCGA  1952

 484 euThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThr  500
     :::::||||||||||||| |||||||||   :::||||||||||||||||||
1953 TGGCCTGCTCCTCCGGA...AGTTTCATCTACTGGTGGGTGCCGGAGACC  1999

 501 LysGlyGlnSerLeuAlaGlu  507
     ||||||:::: |||:::: ||
2000 AAGGGGAAGGCTAACGCCCGAC  2020
```

alignment_scores:
       Quality: 284.50   Length: 581
       Ratio: 1.078   Gaps: 16
Percent Similarity: 45.439   Percent Identity: 20.138

alignment_block:
US-09-652-292-2 x US-09-031-392-1 ..
Align seg 1/1 to: US-09-031-392-1 from: 1 to: 2343

APPLICANT: Tartaglia, Louis A.
APPLICANT: Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Fish & Richardson P.C.
    STREET: 225 Franklin Street
    CITY: Boston
    STATE: MA
    COUNTRY: USA
    ZIP: 02110-2804

COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette
    COMPUTER: IBM Compatible
    OPERATING SYSTEM: Windows95
    SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/299,549
    FILING DATE: 26-APR-1999
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 09/031,392
    FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
    NAME: Meiklejohn, Ph.D., Anita L.
    REGISTRATION NUMBER: 35,283
    REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617/542-5070
    TELEFAX: 617/542-8906
    TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 2343 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    FEATURE:
        NAME/KEY: Coding Sequence
        LOCATION: 73...1761
US-09-299-549-1

```
alignment_scores:
            Quality:  284.50                      Length:   581
            Ratio:      1.078                      Gaps:      16
Percent Similarity:   45.439      Percent Identity:  20.138

alignment_block:
US-09-652-292-2 x US-09-299-549-1

Align seg 1/1   to: US-09-299-549-1   from: 1   to: 2343

  10 LeuCysAlaSerValSerLeuLeuGlyLeuGlyGlyLeuThrPhe......GlyTy  24
     |||  |||||||||||::|::                         |||  ||||
 241 CTCTGGCCTCCCTCGCGGCGGCGCCTTCGGCGTCCTCCTTCCTCCTCACGGGCTA  290

  24 rGluLeuAlaValIleSerGly........................           31
     |::|||:|||:::|||::                                        
 291 CAACCTGTCGGTGGTGAATGCCCCCACCCGTACATCAAGGCCTTTACA          340

  32 ......AlaLeuLeuProLeuGlnLeu                                38
                         |||  |||
 341 ATGAGTCATGGGAAAGAAGGCATGACGTCCAATAGACCCAGACACTCTG         390

  39 AspPheGlyLeuSerCysLeuGluGluGlnGlnGluValGlySerLeuLe        55
     :::    |||  :::||    ||:::||||||||
 391 ACTCTGCTCTGGTCTGTGACTGTGTCCATATTCGCCATCGGTGGA...CT        437

  55 uLeuGlyAlaLeuLeuAlaSerLeuValGlyPheLeuIleAspCysT           72
     |:::|||::::|||||::::::::                  |||
 438 TGTGGGACGTTAATTGTGAAGATGATTGAAAGGTTCTT......              477
```

```
1048 TGCTACCAGCTCTGTGGCCTCAATGCAATTTGGTTCTATACCAACAGCAT 1097
         ||||||   |||                   |||:::::::::||
 257 ePheSerSerValGlyPheHisGlyGlySerSerAlaValLeuAlaSerV  274
     ||||::::::||:::                     :::::
1098 CTTTGGAAAAGCTGGCATC...CCTCCGCAAAGATCCCATACGTCACCT  1144

 274 aIGlyLeuGlyAlaValLysValAlaAlaThrLeuThrAlaMetGlyLeu  290
     :::::  |||:::::::::   |||:::::     ::   :::
1145 TGAGTACAGGGGGCATCGAGACTTTGGCTGCCGTCTTCTCTGGTTTGGTC 1194

 291 ValAspArgAlaGlyArgAlaLeuLeuLeuAlaGlyLyCysAlaLeuMe  307
     ::::::|||  |||||||::::||  ||||||:::|| :::||||
1195 ATTGAGCACCTGGGACGGGAGCACCCCTCCATTGGTGGCTTTGGGCTCAT 1244

 307 tAlaLeuSerValSerGlyIleGlyLeuValSerPheAlaValProMetA  324
     |::::|||                  :||||
1245 GGGCCTCTTC......TTTGGGACCCTCACCATCACGCTG......     1278

 324 spSerGlyProSerCysLeuAlaValProAsnAlaThrGlyGlnThrGly  340
     ......                                            1278

 341 LeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeuProProIlePr  357
     ||||||||||                      :::  |||   ::||
1279 ............ACCCTGCAGGACCACGCC...CCCTGGGTCCC        1307

 357 oArgThrAsnGluAspGlnArgGluProLeuLeuSerThrAlaLysLysT  374
     ......                                            1308

1308 C......                                            1308

 374 hrLysProHisProArgGlyAspProSerAlaProProArgLeuAla    390
     ......                                            1308

 391 LeuSerSerAlaLeuProGlyProProLeuProProAlaArgGlyHisAlaLe 407
     ......                                            1308

 407 uLeuArgTrpThrAlaLeuLeuCysLeuMetValPheValSerAlaPheS  424
     ::  ::::::::::  ||  |||::::::::||||
1309 ......TACCTGAGTATCGTGGGCATTCTGGCCATCATCGCCTCTTTCT   1351

 424 erPheGlyPheProValThrTrpLeuValLeuSerGluIleTyrPro     440
     |||  |||              :::  |||::::::||||:::
1352 GCAGTGGCCAGGTGGCATCCCGTTCATCTTGACTGTGAGTTCTTCCAG   1401

 441 ValGluIleArgGlyArgAlaPheAlaPheCysAsnSerPheAsnTrpAl  457
     :::  |||||||  ||||||             :::  ||||||
1402 CAATCTCAGCGGCCGGCCGCTGCCTTCATCATTGCAGGCACCGTCAACTGGCT 1451

 457 aAlaAsnLeuPheIleSerSerPhe............LeuA          468
     :::||:::  :::|||                    ||||
1452 CTCCAACTTGCTGTTGGGGCTCCTCTCTCCCATTCATTCAGAAAAGTCTGG 1501

 468 sGlyIleGlyThrIleGlyLeuSerTrpPheLeuThrPheLeuGlyLeu  484
     ||                            :::  |||::::::||||
1502 AC...... ACCTACTGTTTCCTAGTCTTTGCTACA                1530

 485 ThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLy  501
     :::::                  :::||||||||||||
1531 ATTTGTATCACAGGTGCTATCTACCTGTATTTTGTGCTGCCTGAGACCAA 1580

 501 sGlyGlnSerLeuAlaGluIleAspGlnGlnInPheGlnLysArg      515
     ||||    ||||||:::||  |||   ::||||||||
1581 AAACAGAACCTATGCAGAAATCAGCCAGGCATTTTCCAAAAGG         1623
```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-299-549-1

```
 72 yrGlyArgLysGlnAlaIleLeuGlySerAsnLeuLeuAlaGly  88
    ||||:|||| ::::::::|||::::::|||
478 .GGGAGGAAGCACACTTTGCTGGCCAATAATGGGTTTGCAATTTCTGCT 525

 89 SerLeuThrLeuGlyLeu...... AlaGlySerLeuAlaTrpLeuVa  102
    :::|||            |||:||||  |||::
526 GCAATTGCTGATGGCCTGCTCCGCTCCAGGCAGGAGCCTTTGAAATGCTCAT 575

102 lLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerMetAlaC  119
    :::|||||| :::::::|||::    :::::::
576 TGTTGGGACGCTTCATCATGGGCATAGATGGGAGCCGTCGCCCTCAGTGTGC 625

119 ysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeu  135
    :::|||||| ::|||||::    ||| |||
626 TCCCCATGTACCTCAGTGAGATCTCACCCAAGGAGATCCGTGGCTCTCTG 675

136 ValSerLeuTyrGluAlaGlyIleLeuThrValGlyIleLeuLeuSerTyrAl  152
    :::|||||::    :::|||||::  ::
676 GGGCAGGTGACTGCCATCTTTATCTGCATTGGCGTGTTCACTGGGCAGCT 725

152 aLeuAsnTyr...AlaLeuAlaGlyThrProTrpGlyTrpArgHisMetP  168
    :|||:::               |||  |||         :::::::|
726 TCTGGGCCTGCCCCGAGCTGCTGGAAAGGAGAGTACCTGGCCATACCTGT 775

168 heGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuPhe  184
    |||||    :::|||||||||||  |||||||||  |||
776 TTGGAGTGATTGTGGTCCCTGCCGTTGTCCAGCTGCTGAGCCCTTCCCTTT 825

185 LeuProAlaGlyThrAspGluThrAlaThrHisLysAspLeuIleProLe  201
    |||                       ||| |||||
826 CTC... 828

201 uGlnGlyGlyGluAlaProLysLeuGlyProGlyArgProArgTyr......  216
    |||:||||:::       |||||||||
829 ......... CCGGACAGCCCACGCTACCTGC 850

216 ..........  216
......... 216

216 ..........  216
TCTTGGAGAAGCACAACGAGGCAAGAGCTGTGAAAGCCTTCCAAACGTTC 900

216 ..........  216
......... 216

217 ...................... SerPheLeuAspPheArgA  224
                            ||| ||:::::|||:::||
901 TTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGGTCCTGGCTGAGAG 950

224 laArgAspAsnMetArgGlyArgThrThrValGlyLeuGlyLeuValLeu  240
    :::|||:::|||||::||||::
951 CCACGTGCAGGAGGAGCATCCCGCCTGGTGTCCGGTCCGAGCTGCTGAGAG 1000

241 PheGlnLeuThrGlyGlnLeuProAsnValLeuCysTyrAlaSerThrI  257
    |||||| |||           |||||||:::
1001 CT...CCCTACGTCCCGCTGGCAGGTGGTCACCCTGATTGTCACCATGGCC 1047

257 ePheSerSerValGlyPheHisGlyGlyGlySerSerAlaValLeuAlaSerV  274
    ||||:::::|||         :::::: :::
1048 TGCTACCAGCTCTGTGGCCTCAATGCAATTTGGTTCTATACCAACAGCAT 1097

274 alGlyLeuGlyAlaValLysValValAlaAlaAlaThrLeuThrAlaMetGlyLeu  290
    :::: |||         |||:::  |||    :::
1098 CTTTGGAAAAGCTGGGATC...CCTCCGGCAAAGATCCCATACGTCACCT 1144

291 ValAspArgAlaGlyArgArgAlaLeuLeuLeuGlyLeuAlaGlyCysAlaAlaLeuMe  307
    :::||||| ::|||:::|||||||:::||| :::||||:
1145 TGAGTACAGGGGCATCGAGACTTTGGCTCCGTCTTCCTGGTTTGGTC 1194

307 tAlaLeuSerValSerGlyIleLeuGlyLeuValSerPheAlaValProMetA  324
    |:::|||        :::|||::||||
1195 ATTGAGCACCTGGGACGGAGACCCTCCTCATTGGCTGTTTGGGCTCAT 1244

324 spSerGlyProSerCysLeuAlaValAlaValProAsnAlaThrGlyGlnThrGly  340
    |:::|||
1245 GGGCCTCTTC...... TTTGGGACCCTCACCATCACGCTG...... 1278
```

```
324 spSerGlyProSerCysLeuAlaValAlaValProAsnAlaThrGlyGlnThrGly  340
     ::
1278 ................................................... 1278

341 LeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeuProProIlePr  357
    |||||||| :::      |||  |||
1279 ............... ACCCTGCAGGACCACCCC...CCCTGGGTCCC 1307

357 oArgThrAsnGluAspGlnArgGluProIleLeuSerThrAlaAlaLysLysT  374
    |
1308 C................................................. 1308

374 hrLysProHisProArgSerGlyAspProSerAlaProProArgLeuAla  390
    
1308 ................................................. 1308

391 LeuSerSerAlaLeuProGlyProProLeuProProAlaArgGlyHisAlaLe  407
    
1308 ................................................. 1308

407 uLeuArgTrpThrAlaLeuLeuCysLeuMetValPheValSerAlaPheS  424
    ::      :::::::|||       ||:::
1309 ......TACCTGAGTATCGTGGGCATTCGGCCATCATCGCCCTCTTTCT 1351

424 erPheGlyPheGlyProValThrProLeuValLeuSerGluIleTyrPro  440
    ::       :::|||::       ||:::||
1352 GCAGTGGGCCAGGTGGCATCCCGTTCATCTTGACTGGTGAGTCTTCCAG 1401

441 ValGluIleArgGlyArgAlaPheAlaPheCysAsnSerPheAsnTrpAl  457
    :::  ::       |||||      :::|||||
1402 CAATCTCAGCGGCCGCGGCTGCCTTCATCATTGCAGGCACCGTCAACTGGCT 1451

457 aAlaAsnLeuPheIleSerSerLeuSerPhe.............. LeuA  468
    ::||||::  :::::::||| |||            |||
1452 CTCCAACTTTGCTGTTGGGCTCCTCTTCCCATTCATTCAGAAAAGTCTGG 1501

468 spLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuTyrGlyLeu  484
    ||                   |||||||:::::::|||
1502 AC.................. ACCTACTGTTTCCTAGTCTTTGCTACA 1530

485 ThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLy  501
    |||               |||||||||||||||
1531 ATTTGTATCACAGGTGCTATCTACCTGTATTTGTGCTGCTGAGACCAA 1580

501 sGlyGlnSerLeuAlaGluIleAspGlnLeuPheGlnLysArg  515
    |:::::::||||:::|||||||:::||| ||:::::||
1581 AAACAGAACCTATGCAGAAATCAGCCAGGCATTTTCCAAAAGG 1623
```

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-501-572-6

seq_documentation_block:
; Sequence 6, Application US/08501572
; Patent No. 6063623
; GENERAL INFORMATION:
;   APPLICANT: Koepsell, Hermann
;   APPLICANT: Grundeman, Dirk
;   APPLICANT: Gorboulev, Valentin
;   TITLE OF INVENTION: Transport protein which Effects The
;   TITLE OF INVENTION: Transport Of Cationic Xenobiotics and\or Pharmaceuticals,
;   TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Finnegan,Henderson,Farabow,Garrett & Dunner
;   STREET: 1300 I Street, N.W., Suite 700
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20005-3315
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30

```
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/501,572
;     FILING DATE:
;     CLASSIFICATION:  424
;  ATTORNEY/AGENT INFORMATION:
;     NAME:  Toohey, Kimberlin M
;     REGISTRATION NUMBER:  35,391
;  REFERENCE/DOCKET NUMBER:  02481.1453-00000
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (202)408-4000
;     TELEFAX:  (202)408-4400
;  INFORMATION FOR SEQ ID NO:  6:
;  SEQUENCE CHARACTERISTICS:
;     LENGTH:  1896 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-501-572-6

alignment_scores:
     Quality:          226.50          Length:          570
     Ratio:            0.956           Gaps:            22
Percent Similarity:    41.579   Percent Identity:   20.175

alignment_block:
US-09-652-292-2 x US-08-501-572-6   .

Algn seg 1/1  to: US-08-501-572-6   from: 1  to: 1896

39  AspPheGlyLeuSerCysLeuGluGln.......GluPheLeuValGl   52
    ::||||::||||  |||              :::::::|||
556 GAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTC  605

52  yserLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyPheLeuI   69
    ::||||  ::||||    ::::::::||||||    |||::::::
606 ATCAGTGAATGTAGGATTCTTTATTGGCTCTGATGAGTATCGGCTACATAG  655

69  leAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeu   85
    ||::||||  |||::|||||     |||:::||    :::|||:::
656 CAGACAGGTTTGGCCGGTAAGCTCTGCCTCCTAACTACAGTCCTCATAAAT  705

86  LeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuVa  102
    |||::::::::    :::::::||::||||||||  |||  :::|||
706 GCTGCAGCTGGAGTTCTCATGGCCATTTCCCAACCTATACGTGGATGTT  755

102 lLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAlaC  119
    ::::  ||| ||  :|| ||||||  ::: :: :::
756 AATTTTTCGCTTAATCCAAGGACTGGTCACAAAGCAGGCTGGTTAATAG  805

119 ysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeu  135
    ::::::::::::|||::::::|||||| |||  ||| :::|||
806 GCTACACTCCTGATTACAGAATTGTTGGGGGAGATATCGGAGAACAGTG  855

136 ValSerLeuTyrGluAlaGlyIleThrArgGlyIleLeuLeuSerTyrAl  152
    :::||||  |||||||||::::::::::||::::|
856 GGGATTTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTCTAGCTGG  905

152 aLeuAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpProArgHis....  166
    :::::  ||||||||||  ||| ||| :::||| :::
906 GGTGGCTTACGCACTT.....CCTCACTGGAGTGGTTGCAGTTCACAG  949

167 .................MetPheGlyTrpAla..........  171
                      ::::: |||::
950 TTGCTCTGCCCAACTTCTTCTTCTCTATTACTGGTCCATACCTGAG  999

171 .................ThrAlaProAlaValLeuG  178
                     ::: |||||| ||||
1000 TCTCCCAGTGGCTGGCTCGATCTCCCAGAATAAGAATGCTGAAGCCATGAGAAT  1049

172 .................ThrAlaProAlaValLeuG  178
                     ::: |||||| ||||
```

```
1050 CATTAAGCAGACATCGCAAAGAAAAATGAAAATCTCTACCGCCTCCCTTC  1099
178 lnSerLeuSerLeuPheLeuProAlaGlyThrAspGluThrAlaThr  194
    ||  |||  |||              |||
1100 AGCGGCCTGAGACTT.................GAAGAGGAAACT  1125
195 HisLysAspLeuProLeuGlnGlyGlyGluAlaProLysLeuGlyPr  211
    ||  |||                  |||
1126 GGCAAGAAATTGAACCCT.........  1143
211 oGlyArgProArgTyrSerPheLeuAspPheAlaArgAlaArgAspAsnM  228
    |||||||||||  ||::
1144 ..........TCATTTCTTGACTTGGTCAGAACT...CCTCAGA  1174
228 etArgGlyArgThrThrValGlyLeu  236
    ::|||  :::|||    :::
1175 TAAGGAAACATACTATTGAATTGATGTACAACTGGTCACGAGCTCTGTG  1224
237 ..........GlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnVa  250
              |||||||:::  ::  |||:::|||:::    |||:::
1225 CTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGAC...AATAT  1271
250 lLeuCysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlyS  267
    |||  :::    ::::::|||:::::
1272 C....TACCTCGATTTCTTCTACTCTGCCCTG.................  1299
267 erSerAlaValLeuAlaSerValGlyLeuGlyAlaLysValAlaAlaAla  283
                                          ||| |||
1300 ...................GTTGAATTCCCAGCT  1314
284 ThrLeuThrAlaMetGlyLeuValAspAlaArgAlaGlyAlaArgAlaLeuLe  300
    :::::  ::  :::::::|||||| |||||||
1315 GCCTTCATGATCATCCTCATTATCGACCGCCATCGGACGCCGTTACCCTTG  1364
300 uLeuAlaGlyCysAlaLeuMetAlaLeuServalSerGlyIleGlyLeuV  317
    ::::::                    :::::::::
1365 GGCTGCATCA.................  1374
317 alSerPheAlaValProMetAspSerGlyProSerCysLeuAlaValPro  333
    |||::::::
1375 ..........AATATGTTGCAGGGGCAGCCTGTCTGGCC.......  1404
334 AsnAlaThrGlyLeuGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAs  350
    ::::::  :::::::::||||||||||
1405 ..........TCAGTTTTTTATACCTGGTGAT.........  1425
350 pSerSerLeuProProIleProArgThrAsnGluAspGlnArgGluProI  367
    ...................................  1425
367 leLeuSerThrAlaLysThrLysProHisProArgSerGlyAspPro  383
    ...................................  1425
384 SerAlaProProArgLeuAlaLeuSerAlaLeuProGlyProProLe  400
    ...................................  1425
400 uProAlaArgArgHisAlaAlaLeuArgTrpThrAlaLeuLeu.....C  415
                                ||::|||
1426 CTACAATGGCTAAAAATTATTATCTCAT  1453
415 ysLeu.....MetValPheValSerAlaPheSerPheGlyPro  429
    ||  |||||          ::::::|||
1454 GCTTGGGAAGAATGGGGATCACAATGGCCTATGAG.........  1488
430 ValThrTrpLeuValLeuSerGluIleTyrProValGluIleArgGlyAr  446
    :||||| |||||  :::::::::
1489 ATAGTCTGCCTGGTCAAGTGCTGAGCTGTACCCCACATTCATTAGGAAТCT  1538
446 gAlaPheAlaPheCysAsnSerPheAsnTrpAlaAlaAsnLeuPheIleS  463
    ::: ||:::
1539 TGGCGTCCACATCTGTTCC.................  1557
```

```
                    Ratio:  0.956              Gaps:           22
Percent Similarity: 41.579    Percent identity: 20.175

alignment_block:
US-09-652-292-2 x US-09-040-444-6

Align seq 1/1 to: US-09-040-444-6    from: 1   to: 1896

  39 AspPheGlyLeuSerCysLeuGluGln.........GluPheLeuValGl  52
     ::!:!!!  !!!   !!!!    :!!
 556 GAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTC 605

  52 ySerLeuLeuGluGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuI  69
     !!!!:!!  ::!!!!!!!:!!!!!!:!!!!  !!!!!!!!!!!:!!!!!!
 606 ATCAGTGAATGTAGGATTCTTTATTGGCTCATGAGTATCGGCTACTACATAG 655

  69 leAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeu  85
     !!! :!!!!!!!!!!!!  :: !!!:
 656 CAGACAGGTTTGGCGTAAGCTCTGCCTCCTAACTACAGTCCTCATAAAT  705

  86 LeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuVa 102
     :!!:!!!!!!!!:!!!!! !!::::!! :::!!
 706 GCTGCAGCTGGAGTTCTCATGGCCATTCCCCAACCTATACGTGGATGTT  755

 102 lLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAlaC 119
     ::!  !!! !!!! :!!!::!!! !!!:!! !!::
 756 AATTTTCGCTTAATCCAAGGACGTGGTCAGCAAAGCAGGCTGGTTAATAG 805

 119 ysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeu 135
     :::!!!:::::::  !!!!!!!!!:!!:!!! :!!  :!!!
 806 GCTACATCCTGATTACAGAATTTGTTGGGGGAGATATCGGAGAACAGTG  855

 136 ValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAl 152
     :::!!!!!!!!!!!  !!!!!!::! :!!!!!! !!!:!!:
 856 GGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCTTGGTGCTAGCTGG 905

 152 aLeuAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpArgHis........ 166
     ::!!!!!!!!!  !!!:!! !!! !!!  !!!
 906 GGTGGCTTACGCACTT......CCTCACTGGAGGTGGTTCAGTTCACAG  949

 167 ..................................MetPheGlyTrpAla.. 171
                                       :::!!!  :!!!::
 950 TTGCTCTGCCCAACTTCTTCTTCTTGCTCTATTACTGGTGCATACCTGAG 999

 171 .................................................. 171
 172 ...............................ThrAlaProAlaValLeuG 178
                                    :::  !!!!!! !!!!!
1000 TCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTGAAGCCATGAGAAT 1049
1050 CATTAAGCACATCGCAAAGAAAATGGAAAATCTTACCCGCCTCCCTTC   1099

 178 lnSerLeuSerLeuPheLeuProAlaGlyThrAspGluThrAlaThr     194
     !! !!!  !!!  !!    !!!!!:!!: !!!
1100 AGGCCCTGAGACTT...............GAAGAGGAAAACT         1125

 195 HisLysAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeuGlyPr 211
     !!!!!!!!!:!! :::!!:!!:!!  !!!!!:!!!!!:
1126 GGCAAGAAATTGAACCCT................................ 1143

 211 oGlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAspAsnM 228
     :::!!!:!!!!!!!!!!!! !!!:  !!
1144 ................TCATTTCTTGACTTGGTCAGAACT..CCTCAGA   1174

 228 etArgGlyArgThrThrValGlyLeu......................... 236
     ::!!!  :::!!!  :!!
1175 TAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCTCTGTG 1224

 237 .........GlyLeuValLeuPheGlnIleThrGlnLeuGlyProAsnVa 250
              !!:::!!!:::!!:!!  :!:!!!:::!! !:!!:
1225 CTCTACCAGGGCCTCATCATGCACATGGCCGTTGCAGGTGAC...AATAT 1271

 250 lLeuCysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlyS 267
```

```
 463 erLeuSerPheLeuAspLeuIleGlyThrIleGlyLeuSerTrpThrPhe 479
     !!:::   !!! !!!  :!!
1558 ....TCAATGTGTGACATTGGTGGCATCATC.........ACGCCATTC  1593

 480 LeuLeuTyrGlyLeuThr............................... 485
     !!:!!  !!!!!!!
1594 CTGGTCTCACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTT 1643

 486 .AlaValLeuGlyLeu.....GlyPheIleTyrLeuPheValProGluT  500
      ::!!!!!!!!!!!!     !!  ::: !!!:::::!!!!!!
1644 CGGCCGTACTTGGCTTCGTTGCTGGAGGTCGTCGTGCTGTTGCTTCCAGAAA 1693

 500 hrLysGlyGlnInSerLeuAlaGluIleAspGlnInPheGlnLysArgArg 516
     !!!!!!:!:::::!! !!!  :::::! :::
1694 CTAAAGGAAAGCTTTGCCTGAGACCATCGAGGAAGCCGAAAATATGCAA   1743

 517 PheThrLeuSerPheGlyHisArgGlnAsnSerThrGlyIleProTyrSe 533
     !!!!!!!!!!!!!:!!!!!
1744 .........AGACCAAGAAAAAAATAAAGAAAAGATGATTTACCT       1778

 533 rArgIleGlu 536

1779 CCAAGTTCAG 1788
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-040-444-6

alignment_scores:
    Quality: 226.50   Length: 570

1779 CCAAGTTCAG 1788

seq_documentation_block:
; Sequence 4, Application US/08501572
; Patent No. 6063623
; GENERAL INFORMATION:
;   APPLICANT: Koepsell, Hermann
;   APPLICANT: Grundeman, Dirk
;   APPLICANT: Gorboulev, Valentin
;   TITLE OF INVENTION: Transport protein Which Effects The
;   TITLE OF INVENTION: Transport Of Cationic Xenobiotics and\or Pharmaceuticals,
;   TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Finnegan,Henderson,Farabow,Garrett & Dunner
;   STREET: 1300 I Street, N.W., Suite 700
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20005-3315
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/501,572
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   CLASSIFICATION: 424
;   NAME: Toohey, Kimberlin M
;   REGISTRATION NUMBER: 35,391
;   REFERENCE/DOCKET NUMBER: 02481.1453-00000
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)408-4000
;   TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1882 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-501-572-4

alignment_scores:
       Quality: 201.00   Length: 546
       Ratio: 0.824   Gaps: 24
Percent Similarity: 44.689   Percent Identity: 23.260

alignment_block:
US-09-652-292-2 x US-08-501-572-4

Align seg 1/1 to: US-08-501-572-4   from: 1   to: 1882

    37 GlnLeuAspPheGlyLeuSerCysLeuGluGlnGluPheLeuValGlySe 53
       :::::|:||::: |||| :::|:::|:|:|:::
   479 AAAGTGGACCTTTTCAGTCCTGTGTGAAC.................... 508

    53 rLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleA 70
       ::|||::||:::||||| :::|||||||| |||:::|| |
   509 .......TTGGGCTTCTTCCTGGGCTCCCTGGTTGTGGGTTACATTGCAG 551

    70 spCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu 86
       || |||||||| :||| :::|||:::|||
   552 ACAGGTTTGGCCGTAAGCTCTGTCTCTTGGTGACCACGCTGGTCACATCT 601

    87 AlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValLe 103
       :::::||||| ||| ::: ||:::|||| :::::||
   602 GTGTCCGGTGTGCTAACAGCGGTGGCCCCAGACTATACATCCATGTTGCT 651

103 uGlyArgAlaValValGlyPheAlaIleSerSerMetAlaCysC 120
    | ||| ::: ||| ||| | ||||
652 CTTTCGCCTGCTGCGAGGGCATGGTCAGCAAGGGCAGCTGGGTGTCCGGCT 701

120 ysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuVal 136
    ||::: :::||| ||||| ||| :: ||||
702 ATACCTTGATCACAGAGTTTGTCGGCTCTGGCTACAGAGAACGACGGCC 751

137 SerLeuTyrGluAlaGlyIleThrValGlyIleIleLeuLeuSerTyrAlaLe 153
    |||||:: :::|:||||||||||||:::::
752 ATTTTGTACCAGATGCCCTTCACAGTGGGCTAGTGGGCTTGGCCGGGGT 801

153 uAsnTyrAlaLeuAlaLeuGlyThrProTrpProGlyTrpArgHisMetPheGlyT 170
    | :||||::|| ||| ||| ||| :||:::
802 GGCCTATGCCATT.....CCAGACTGGCCGCTGGCCTCCAGCTAGCT.... 841

170 rpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheLeuPro 186
    ::: ||| ||| ||| ::|||:::||||||||
842 ..GTGTCCCTGACCTACCTTCCTTCCTGCTGTATTACTGGTTGTCCCA 889

187 AlaGly..............ThrAspGluThrAlaTh 194
    ||| :: |||
890 GAATCCCCCCGGTGGCTGTTGTCCCAGAAGAGAACCACGCGAGCTGTCAG 939

194 rHisLysAspLeuIleProLeuGlnIleGlyGluValAlaPro.......... 207
    ::: |||| ||||:::|||
940 GATAATGGAGCAAATTGCACAGAAGACGGGAAGGTGCCTCCTGCTGACC 989

208 ....LysLeuGlyProGly 212
    ||| :: |||
990 TGAAGATGCTCGCCTTGAGGAGGATGCCTCAGAAAAGCGAAGPCCT... 1036

213 ArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAspAsnMetAr 229
    |||||||:::|||:::||||:::|
1037 .............TCGTTTGCCGACCTGTTCCGGACT...CCCAACCTGAG 1071

229 gGlyArgThrThrValGlyLeuGlyLeu.............ValLeuP 241
    | ::|||| ||| ||| ||||||||
1072 GAAGCACACCGTCATCCTGATGTATCCTGATGGTTCTCTTGTGCTGTCGTGT 1121

241 heGlnGlnLeu........ThrGlyGlnProAsnValLeuCysTyrAla 254
    ::||| ||| :: ||||:::|||
1122 ACCAGGGTCTCATCATGCACGTGGGAGCCACAGGGGCCACAACCTCTACCTG 1171

255 SerThrIlePheSerSerValGlyPheHisGlyGlySerSerAlaValLe 271
    ::: :::||:::|||||:::
1172 GACTTCTTTTATTCTTCTCTG............................. 1192

271 uAlaSerValGlyLeuGlyAlaValLysValAlaAlaThrLeuThrAlaM 288
    ||:||| :::
1193 .........GTGGAATTCCCCGCGGGCCTTCATCATCC 1220

288 etGlyLeuValAspArgArgAlaGlyArg..........ArgAlaLeu 299
    :: :::||||| ||||| ||
1221 TGGTCACCATTGACCGCATTGGCCGCATCTACCCCAATAGCGGGCCTCGAAT 1270

300 LeuLeuAlaGlyCysAlaLeuMetAlaLeu..SerValSerGlyIleGlyLeuGlyL 316
    |||:::|||::: ||| :: :::||:::||
1271 CTGGTGACGGGGGCAGCCTGCCTCCTCATGATCTTTATCCCGCATG..... 1316

316 euValSerPheAlaValProMetAspSerGlyProSerCysLeuAlaVal 332
    ||||||
1317 ....AGCTGC................... 1322

333 ProAsnAlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGl 349
    |||||| :: |||
1323 ........ACTGGTTGAACG................ 1334

349 nAspSerSerLeuProProIleProArgThrAsnGluAsnGlnGlnArgGluP 366
    ||||||
1335 ....TTACCC................TCGC 1344

366 roIleLeuSerThrAlaLysThrLysProHisProArgSerGlyAsp 382
    ||::::|||
1345 CTGTCTTG................................. 1352

383 ProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGlyProPr 399
    |||:::
1353 ..............GCCGTATGGGGGCCACC 1369

399 oLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrArgAlaLeuLeuCysL 416
    ||||
1370 ATTG..............TGCTCAGATGGTCTGC........... 1390

416 euMetValPheValSerAlaPheSerPheGlyPheGlyProValThrTrp 432
    1390

433 LeuValLeuSerGluIleTyrProValGluIleArgGlyArgAlaPheAl 449
    |||||| :::||:::||:::|| :::::
1391 CTGGTGAACGGTGAGCTGTACCCTACATTCATCAGGAATCTTGGGATGAT 1440

449 aPheCysAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerSerLeuSerP 466
    ||:::|| :::::::::|||
1441 GGATTGCTCTGCCCCTGTGACCTGGGTGGGATCTTCACCCCCTTCATGG 1490

466 heLeuAspLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuLeuTyr 482
    ::|||||||||||||::: ::: :::
1491 TGTTCAGGCTGATGGGAAGTT............TGGCAAGCCCTGCCCCTC 1528

483 GlyLeuThrAlaValLeuGlyLeu.......GlyPheIleTyrLeuPheVa 497
    ||| ||||||||||||| ||| | :: ||||||
1529 ATTTTGTTTGGGGTTTGGGCCTGACTGCTGGGGCCCATGACTCTTTCTCT 1578

497 lProGluThrLysGlyGlnSerLeuLeuAlaGlu.Ile 508
    :||||||||||||| :::||| ||| :::
1579 CCCAGAGACCAAGGGTGTGGCTTTGCCTGAGACTATTGAAGAAGCAGAGA 1628

509 ....AspGlnGlnPheGlnLysArgArgPheThrLeuSer 520
    ::||| ||||||||
1629 ACCTGGGAGGAGGAGAAATCAAAGGCCAAGAAAAACACGATTTACC..... 1673

520 rPheGlyHisArgGlnAsnSerThrGlyIleuPro 531
    ||| ||||||:::||| |||
1674 .TTCAGGTCCAAACAGCAAGTCCTCAAGTACCT 1706

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-040-444-4

```
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  O'Connor, Steven P
;       REGISTRATION NUMBER:  41,225
;       REFERENCE/DOCKET NUMBER:  2481.1453-01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)408-4000
;       TELEFAX:   (202)408-4400
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1882 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-040-444-4

alignment_scores:
       Quality:     201.00          Length:      546
       Ratio:         0.824          Gaps:         24
Percent Similarity:   44.689    Percent Identity:  23.260

alignment_block:
US-09-652-292-2 x US-09-040-444-4

Align seg 1/1  to:  US-09-040-444-4  from: 1  to: 1882


  37  GlnLeuAspPheGlyLeuSerCysGluGlyGlnGluPheLeuValGlySe  53
      ::::::|:|||  |:   ::::|||||:|:|||
 479  AAAGTGGACCTTTTCAGTCCTGTGAAC................................. 508

  53  rLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleA  70
      |||||   :::||:|:|||||||||   |||::::::   ::|
 509  .........TTGGGCTTCTCCTGGGCTCCCTGGTTGTGGTTACATTGCAG  551

  70  spCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu  86
      ||::::||||||||  :::::|||    :::::||||||||
 552  ACAGGTTGCCGTAAGCTCTGTCTCTTGGTGACCACGCTGGTCACATCT  601

  87  AlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValIle  103
      ::::::||||::::  ::::|||||||:::::|||
 602  GTGTCCGGTGTGCTAACAGCGGTGGCCCCAGACTATACATCCATGTTGCT  651

 103  uGlyArgAlaValAlaValGlyPheAlaIleSerLeuSerMetAlaCysC  120
      | |||  :::  ||:|:::   ::  |||   :||||
 652  CTTTCGCCTGCTGCAGGGCATGGTCAGCAAGGCAGCTGGGTGTCCGGCT  701

 120  ysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuVal  136
      ::::||:::|||||||||       |||  |||  :::
 702  ATACCTTGATCACAGAGTTTGTCGGCCTCTGGCTACAGGAGAACGACGGCC  751

 137  SerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAlaLe  153
      |||||||   |||  |||     ||:::|||
 752  ATTTTGTACCAGATGGCCTTCACAGTGGGGCTAGTGGGGCTTGCCGGGT  801

 153  uAsnTyrAlaLeuLeuAlaGlyThrProTrpGlyTrpArgHisMetPheGlyT  170
      |||||||:::       |||  |||  :::  :::
 802  GGCCTATGCCATT.......CCAGACTGGCGCGCTGGCTCCCAGCTAGCT....  841

 170  rpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuPheLeuPro  186
      ::::  |||:::  |||   |||   |||:::|||
 842  ..GTGTCCTGCCTACCCTTCCTCCTTCCGTATTACTGGTTTGTCCCA  889

 187  AlaGly.................................ThrAspGluThrAlaTh  194
      :::::::   |||   :::::|||
 890  GAATCCCCCCCGGTCGTTGTTGTCCCAGAAGAGAACCACGCGAGCTGTCAG  939

 194  rHisLysAspLeuIleProLeuGlnGlyGlyGluAlaPro.................  207
      ||::||:::|
 940  GATAATGGAGCAAAATTGCACAGAAGAACGGGAAGGTGCCTCCTGCTGACC  989

 208  ...............................LysLeuGlyProGly  212
                                      ||| :::|||
```

```
 990  TGAAGATGCTGTCGCCTTGAGGAGGATGCCTCAGAAAAGCGAAGTCCT...  1036

 213  ArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAspAsnMetAr  229
      ||||||  |||||||:|||||||||     ||:::||
1037  ........TCGTTGCCGACCTGTTCCGCACT...CCCAACCTGAG  1071

 229  gGlyArgThrValGlyLeuGlyLeu.................ValLeuP  241
      |  :::|||   |||               ||||||
1072  GAAGCACACCGTCATCCTGATGTATCTATGGTTCTTCTTGTCTGTGTGCTGT  1121

 241  heGlnGlnLeu...........ThrGlyGlnProAsnValLeuCysTyrAla  254
      :::|| |||            :::|||         |||
1122  ACCAGGGTCTCATCATGCACGTGGGAGCCACAGGGGCCAACCTCTACCTG  1171

 255  SerThrIlePheSerSerValGlyPheHisGlyGlySerSerAlaValLe  271
      :::  :::::::||||||:::
1172  GACTTCTTTTATTCTTCTCTG.................................  1192

 271  uAlaSerValGlyLeuGlyAlaValLysValAlaAlaAlaThrLeuThrAlaM  288
      |||:::  |||::::::|||
1193  .....................GTGGAATTCCCCGCGGCCTTCATCATCC  1220

 288  etGlyLeuValAspArgAlaGlyArg.................ArgAlaLeu  299
      ::   ::|||||||||
1221  TGGTCACCATTGACCGCCATTGGCCGCATCTACCCCATAGCGCCTCGAAT  1270

 300  LeuLeuAlaGlyCysAlaLeuMetAlaLeu.SerValSerGlyIleGlyL  316
      |||:::||||   ::::::|||:::|||  :::
1271  CTGGTGACGGGGCAGCCTGCCTCCTCATGATCTTTATCCCGCATG.....  1316

 316  euValSerPheAlaValProMetAspSerGlyProSerCysLeuAlaVal  332
      ||||||            ||||||
1317  ........................AGCTGC.........................  1322

 333  ProAsnAlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGl  349
      :::|||||||  :::        |||
1323  .......ACTGGTTGAACG.....................................  1334

 349  nAspSerLeuProProIleProAlaThrArgAsnGluAspGlnArgGluP  366
      :::
1335  ..........TTACCC.....................................TCGC  1344

 366  roIleLeuSerThrArgAlaLysThrLysProHisProArgSerGlyAsp  382
      ||:::|||
1345  CTGTCTTG.........................................AGCTGC.....  1352

 383  ProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGlyProPr  399
      ||||:::  ||||||||
1353  .................................GCCGTATGGGGGCCACC  1369

 399  oLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysL  416
      ||||       :::
1370  ATTG..........................TGCTGCAGATGGTCTGC.........  1390

 416  euMetValPheValSerAlaPheSerPheGlyPheGlyProValThrTrp  432
1390  .........................................................  1390

 433  LeuValLeuSerGluIleTyrProValGluIleArgGlyArgAlaPheAl  449
      ||||| :::::::|:|||:::||:::   ||||||   :::::
1391  CTGGTGACGCTGAGCTGTACCCTACATTCATCAGGAATCTTGGATGAT  1440

 449  aPheCysAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerP  466
      :::::::||:||   :::
1441  GGATTGCTCTGCCCCTGTGACCTGGATCTTCACCCCCTTCATGG  1490

 466  heLeuAspLeuIleGlyThrIleGlyLeuSerTrpTrpProValThrTrp  482
      ::: ||||||::::::|||
1491  TGTTCAGGCTGATGGAAGTT.......TGGCAAGCCCTGCCCCCTC  1528

 483  GlyLeuThrArgAlaValLeuGlyLeu.............GlyPheIleTyrLeuPheVa  497
      ||| :::   ||||||||:::::
1529  ATTTTGTTTGGGGTTTTGGGCCTGACTGCTGGGGCACCATGACTCTTCTTCT  1578
```

Right margin line numbers (protein): 1036, 229, 1071, 241, 1121, 254, 1171, 271, 1192, 288, 1220, 299, 1270, 316, 1316, 332, 1322, 349, 1334, 366, 1344, 382, 1352, 399, 1369, 416, 1390, 432, 1390, 449, 1440, 466, 1490, 482, 1528, 497, 1578

Left column sequence line numbers: 53, 508, 70, 551, 86, 601, 103, 651, 120, 701, 136, 751, 153, 801, 170, 841, 186, 889, 194, 939, 207, 989, 212

```
497 lProGluThrLysGlyGlnSerLeuAlaGlu.Ile.............. 508
    :||||||||||||       :::||  ||| ::|
1579 CCCAGAGACCAAGGGTGTGGCCTTGCCTGAGACTATTGAAGAAGCAGAGA 1628

509 ..........AspGlnPheGlnLysArgArgPheThrLeuSe 520
              :::||  :::|| |||||||||
1629 ACCTGGGAGGAGAAATCAAAGGCCAAAGAAAACACGATTTACC...... 1673

520 rPheGlyHisArgGlnAsnSerThrGlyIlePro 531
    |||  ::||||   |||    ::||||
1674 .TTCAGGTCCAAACAGGCAAGTCCTCAAGTACCT 1706
```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-501-572-5

seq_documentation_block:
; Sequence 5, Application US/08501572
; Patent No. 6063623
GENERAL INFORMATION:
    APPLICANT: Koepsell, Hermann
    APPLICANT: Grundeman, Dirk
    APPLICANT: Gorboulev, Valentin
    TITLE OF INVENTION: Transport protein Which Effects The
    TITLE OF INVENTION: Transport Of Cationic Xenobiotics and\or Pharmaceuticals,
    TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
    NUMBER OF SEQUENCES: 6
    CORRESPONDENCE ADDRESS:
       ADDRESSEE: Finnegan,Henderson,Farabow,Garrett & Dunner
       STREET: 1300 I Street, N.W., Suite 700
       CITY: Washington
       STATE: D.C.
       COUNTRY: USA
       ZIP: 20005-3315
    COMPUTER READABLE FORM:
       MEDIUM TYPE: Floppy disk
       COMPUTER: IBM PC compatible
       OPERATING SYSTEM: PC-DOS/MS-DOS
       SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/501,572
       FILING DATE:
       CLASSIFICATION: 424
    ATTORNEY/AGENT INFORMATION:
       NAME: Toohey, Kimberlin M
       REGISTRATION NUMBER: 35,391
       REFERENCE/DOCKET NUMBER: 02481.1453-00000
    TELECOMMUNICATION INFORMATION:
       TELEPHONE: (202)408-4000
       TELEFAX: (202)408-4400
    INFORMATION FOR SEQ ID NO: 5:
    SEQUENCE CHARACTERISTICS:
       LENGTH: 1885 base pairs
       TYPE: nucleic acid
       STRANDEDNESS: single
       TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-501-572-5

alignment_scores:
    Quality: 198.50   Length: 515
    Ratio: 0.906   Gaps: 17
Percent Similarity: 42.524   Percent Identity: 20.971

alignment_block:
US-09-652-292-2 x US-08-501-572-5
Align seg 1/1 to: US-08-501-572-5   from: 1 to: 1885

```
 37 GlnLeuAspPheGlyLeuSerCysLeuGluGlnGluPheLeuValGlySe 53
    :::||||||  |||||||||||||   |||:||
511 AAGCTGGACCTCTTTCAGTCCTGTTTGAATGCGGCGGCTTCTTTCTTTGGCCTC 560

 53 rLeuLeuLeuGlyAlaAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleA 70
    ||||                                        ::|
561 TCTCGGTGTT.............................GGCTACTTTGCAG 583

 70 spCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu 86
    ||     :::|||||||    :::||||||||||||:::    ||||||
584 ACAGTTTGGCCGTAAGCTGTCTCCTGGAACTGTGGTCGTCAACGCG 633

 87 AlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValLe 103
    :|||    :||||:: |||||     |||||||    ||||:::  ::||
634 GTGTCGGGGTGCTCATGGCCTTCTGCCCAACTACATGTCCATGCTGCT 683

103 uGlyArgAlaValAlaValGlyPheAlaIleSerLeuSerSerMetAlaCysC 120
    | |||   :::  :: ||||||  :: :::  :::   ||||:|
684 CTTCCGCTCCTGCGAGGCCTGGTCAGCAAGGGCAACTGATGGCTGGCT 733

120 ysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuVal 136
    :::|||   |||:::|||||||||   ::|||   :::||
734 ACACCCTAATCACAGAATTGTTGGCTCGGCTCCAGAAGAACGGTGGCG 783

137 SerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAlaLe 153
    :::|| ||||:::     ||:::||||||     |||:::  ::||
784 ATCATGTACCAGATGGCTTCACGGTGGGCTGTGGCGCTTACCCGGGCT 833

153 uAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpArgHis.......... 166
    | ||| ||||||   |||     ||| ||  ||
834 GGCCTACGCCCTG......CCTCACTGGCCGCTGGCTCGCAGCTGGCAGTCT 877

167 .....................MetPheGlyTrpAlaThrAlaProAla 175
                         :: :::   |||
878 CCCTGCCCACCTTCCTTCCTTCCTTGCTCTACTACTGG.............TGT 915

176 ValLeuGlnSerLeuSerLeuPheLeuProAlaGlyThrAspGluTh 192
    ||| :::|||     ||::|     ::|   |||:::
916 GTGCGGGAGTCCCCTCCGGTGGCTGTTATCACAAAAAGAAACACTGAAGC 965

192 rAlaThrHisLysAspLeuIleProLeuGlnGlyGlyGluAlaPro.... 207
    :     :::  ||| |||  :::|||:::|||     |||
966 AATAAAGATAATGGACCACCATCGCTCAAAGAATGGGAAGTTGCCTCCTG 1015

208 ....................................LysLeuGly 210
                                        |||||
1016 CTGATTTAAAGATGCTTTCCCTCGAGAGGATGTCACCGAAAGCTGAGC 1065

211 ProGlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAspAs 227
    |||||    :::|||:: |||    |||||||||||   ::
1066 CCT............TCATTTGCAGACCTGTTCCGCACG...CCGCG 1097

227 nMetArgGlyArgThrThrValGlyLeuGlyLeu...............V 239
    :::||| |||||  ::: |||
1098 CCTGAGGAAGCGCACCTTCATCCTGATGTACCTGTGGTTCACGGACTCTG 1147

239 alLeuPheGlnGlnLeu.......ThrGlyGlnProAsnValLeuCys 252
    |||||||||||||||         ||| |||
1148 TGCTCTATCAGGGCTCATCCTGCACATGGGCGGCCACCAGGGGAACCTC 1197

253 TyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerSerAl 269
    |||     |||||||||:::::
1198 TACCTGGATTTCCTTTACTCCGGTCTG....................... 1224

269 aValLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThrLeuT 286
    :::||:::       |||||||:::
1225 ..............GTCGAAATCCCGGGGGCCTTCA 1246

286 hrAlaMetGlyLeuValAspArgAlaGlyArg................Arg 297
    ||||   :: |||||||||||||::||||
1247 TAGCCCTCATCACCATTGACCGGTGGGCCGCATCTACCCCATGCCCGTG 1296

298 AlaLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIl 314
    :::  ||||||||||||:::||    ::::||
1297 TCAAATTTGTTGGCGGGGCAGCCTGCCCTGCCCGTCATTTTTATCTCA..... 1341

314 eGlyLeuValSerPheAlaValProMetAspSerGlyProSerCysLeuA 331
```

Fri Mar 15 14:21:22 2002

us-09-652-292-2.rni

APPLICATION NUMBER: US/09/040,444
FILING DATE: March 18, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Connor, Steven P
REGISTRATION NUMBER: 41,225
REFERENCE/DOCKET NUMBER: 2481.1453-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1885 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-040-444-5

alignment_scores:
Quality: 198.50   Length: 515
Ratio: 0.906   Gaps: 17
Percent Similarity: 42.524   Percent Identity: 20.971

alignment_block:
US-09-652-292-2 x US-09-040-444-5

Align seg 1/1 to: US-09-040-444-5   from: 1 to: 1885

```
 37 GlnLeuAspPheGlyLeuSerCysLeuGluGlnPheLeuValGlySe 53
    ::::||||||:::                ||:::       |||::
511 AAGCTGGACCTCTTTCAGTCCTGTTTGAATGCGGGCTTCCTTCTTTGGCTC 560

 53 rLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyLyPheLeuIleA 70
    ||||                         |||:::|::||
561 TCTCCGGTGTT..........                GGCTACTTTGCAG 583

 70 spCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu 86
    ||  ::::||||||:|  :::|||||||:::: |||||
584 ACAGGTTTGGCCGTAAGCTGTGTCTCCTGGGAACTGTGCTGGTCAACGCG 633

 87 AlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuLeuAlaTrpLeuValIle 103
    ::::::::::::   ::::    ::    |||||     ||||||
634 GTGTCGGGCGTGCTCATGGCCTTCTCGCCCAACTACATGTCCATGCTGCT 683

103 uGlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAlaCysC 120
    ||||    ||||::::  ::   ::         |||||    ||||||
684 CTTCCGCCTGCTGCAGGGCCTGGTCAGCAAGGGCAACTGGATGGCTGGCT 733

120 ysIleTyrValSerGluValGlyProArgGlnArgGlyValValLeuVal 136
    :::::::::::   |||:::||||   ::||    ::::||||
734 ACACCCTAATCACAGAATTTGTTGGCTCGGGCTCCAAGAGAACGGTGGCG 783

137 SerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAlaLe 153
    ::::||:::    :::||||||||||||||   ||::    :::||
784 ATCATGTACCAGATGGCCTTCACGGTGGGCGTGGCGGCCTTACCGGGCT 833

153 uAsnTyrAlaLeuAlaGlyThrProTrpProGlyTrpArgHis......... 166
    |||||||||||      |||  |||  ::
834 GGCCTACGCCCTG..........CCTCACTGGCCGTGGCTGCAGCTGGCAGTCT 877

167 ..........MetPheGlyTrpAlaThrAlaProAla 175
                ||||
878 CCCTGCCCACCTTCCTCTTCCTGCTCTACTACTG..........TGT 915

176 ValLeuGlnSerLeuSerLeuLeuProLeuProAlaGlyThrAspGluTh 192
    |||  ::::|||        :::|||   |||::::
916 GTGCCCGGAGTCCCCTCGGTGGCTGTTATCACAAAAAGAAACACTGAAGC 965

192 rAlaThrHisLysAspLeuProLeuGlnLysGlyGlyGluAlaPro..... 207
    |||        |||   :::::||||||   |||
966 AATAAAGATAATGGACCACATCGCTCAAAAGAATGGGAAGTTGCCTCCTG 1015
```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-040-444-5

```
                                                        1341
331 laValProAsnAlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeu 347
    ||:::
1342 ....CCTGAC..........                               1347
348 LeuGlnAspSerSerLeuProProIleProArgThrAsnGluAspGlnAr 364
                                                        1347
364 gGluProIleLeuSerThrAlaLysLysProHisProArgSerG 381
                                                        1347
381 lyAspProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGly 397
                                                        1347
398 ProProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLe 414
                          CTGCACTGGTTAAACATCAT 1367
414 uCysLeuMetValPheValSerAlaPheSerGlyPheGlyProValT 431
    : :::     |||                  ::
1368 AATCATGTGTGTTGGCCGAATGGAATCACCATTGCAAATACAAATGATCT 1417
431 hrTrpLeuValLeuSerGluIleTyrProValGluIleArgGlyArgAla 447
     ||||||
1418 GC...CTGGTGAATGCTGAGCTGTACCCCACATTCGTCAGGAACCTCAGA 1464
448 PheAlaPheCysAsnSerPheAsnTrpAlaAlaAsnLeu..........Ph 461
    ||:::|||:::        :::::::::
1465 GTGATGGTGTGTTCCTCCGTGTGACATAGGTGGAATAATCACCCCCTT 1514
461 eIleSerLeuSerPheLeuAspLeuIleGlyThrIleGlyLeuSerTrpT 478
    ||||          :::     :::|||
1515 CATAGTCTTCAGGCTGAGGGAGGTC..........          TGGC 1543
478 hrPheLeuLeuTyrGlyLeuThrGlyLeuThrAlaLeuValLeuGly... 492
    |||            |||   ||||||||||     ::
1544 AAGCCTTGCCCCTCATTTTGTTGCGGTGTTGGGCCTTGCTTGCCGCGGGA 1593
493 IleTyrLeuPheValProGluThrLysGlyGlnSerLeuLeuAlaGlu 507
    :::  ||||||||||||||||||:::  |||
1594 GTGACGCTACTTCTTCCAGAGACCAGGGGACGCGTTTGCCAGAG 1638
```

```
478 hrPheLeuTrpTyrGlyLeuThrArgAlaValLeuGlyLeu......GlyPhe 492
    |||      |||||||||||||||||||||||      :::
1544 AAGCCTTGCCCTCATTTGTTGCGGTGTTGGCCTGTTGCCGCGGGA 1593

493 IleTyrLeuPheValProGluThrLysGlyGlnSerLeuAlaGlu 507
    :::    |||||||||||||||||||||||||::::::||  ::
1594 GTGACGCTACTTCTTCCAGAGACCAAGGGGACGCTTTCCCAGAG 1638

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-2

seq_documentation_block:
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;    APPLICANT: FLEISCHMAN, Robert D.
;    APPLICANT: WHITE, Owen R.
;    APPLICANT: FRASER, Claire M.
;    APPLICANT: VENTER, John C.
;    TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;    TITLE OF INVENTION: TUBERCULOSIS
;    FILE REFERENCE: 24366-20007.00
;    CURRENT APPLICATION NUMBER: US/09/103,840A
;    CURRENT FILING DATE: 1998-06-24
;    NUMBER OF SEQ ID NOS: 2
;    SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;    LENGTH: 4403765
;    TYPE: DNA
;    ORGANISM: Mycobacterium tuberculosis
;    FEATURE:
;    OTHER INFORMATION: CDC 1551
;    OTHER INFORMATION: "n" bases at various positions throughout the sequence
;    OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

alignment_scores:
    Quality:        195.00              Length:      517
    Ratio:          0.786               Gaps:        21
Percent Similarity:  47.969      Percent Identity:  22.824

alignment_block:
US-09-652-292-2 x US-09-103-840A-2/rev

Align seg 1/1  to reverse of: US-09-103-840A-2   from: 1  to: 4403765

      12 AlaSerValSerLeuGlyGlyLeuThrPheGlyTyrGluLeuAlaVa 28
         |||::|||      :::||||||||:::::      :::::||
  880438 GCCGAGTGTGTCTGGCCTCGGTGATGGCACATGTGACGTCACCGT 880389

      28 IleSerGlyAlaLeuProLeuGlnLeuAspPheGlyLeuSerCysL 45
         :::||  |||      ||||||||::      |||||||
  880388 GGTCAGCGTCGCGCAACGCACCTTCGTTGCCGATTTCGGG........ 880349

      45 euGluGlnGlnPheLeuValGlySerLeuLeuGlyAlaLeuLeuAla 61
         :::    :::||||      :::    ||||||      ::||||||
  880348 ..TCCACCCAGGCCCGTCGTGGCCTGGACCATGACCGGCTACATGCTCGCG 880301

      62 SerLeu.........ValGlyGlyPheLeuIleAspPCysTyrGlyAr 74
         :::||:::          |||      ||||      ::|||
  880300 TTGGCCACTGTGATCCCGACAGCCGGTTGGGCGGCCGGTTCGGCAC 880251

      74 gLysGlnAlaIleLeuGlySerAsnLeuValLeuAlaGlySerLeuT 91
         :::::::||      ::||||||      |||||||||
  880250 CAGGCGGCTTTTCATGGGTTCGGTATTGGCGTTCACCCTCGGTCACTGC 880201

      91 hrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValLeuGlyArgAlaVal 107
         |||::|||      :::::::::          ||
  880200 TGTCGGCAGTGAGCACCAAACATCTTGCTGCTCATCATATTTCGTGTTGTC 880151

     108 ValGlyPheAlaIleSerLeuSerSerMetAlaCysIleTyrValSe 124
         |||||||      :::      :::
  880150 CAGGGTTTCGGTGGGGCATGCTGACCGCGGTGTCCTTTGCCATCTTGGC 880101
```

```
208 ......................................LysLeuGly 210
                                            |||||||
1016 CTGATTTAAAGATGCTTCCCTCGAAGAGGATGTCACCGAAAAGCTGAGC 1065

211 ProGlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAspAs 227
    |||      |||||   ||||||||||||::
1066 CCT......TCATTTGCAGACCTGTTCCGCACG...CCGCG 1097

227 nMetArgGlyArgThrThrValGlyLeuGlyLeu...............V 239
    :::||||  |||||||  ::: :::|||                      }
1098 CCTGAGGAAGCGCACCTTCATCCTGATGTACCTGTGGTTCACGGACTCTG 1147

239 aLleuPheGlnGlnLeu.........ThrGlyLnProAsnValLeuCys 252
    ||||||:::|||  |||      :::
1148 TGCCTCATCCAGGGGCTCATCCTGCACATGGGCGGCCACCAGGCGGAACCTC 1197

253 TyrAlaSerThrIlePheSerValGlyPheHisGlyGlySerSerAl 269
    |||  :::||||||||||||||||:::::
1198 TACCTGGATTTCCTTTACTCCGCTCTG...................... 1224

269 aValLeuAlaSerValGlyLeuGlyAlaValAlaLysValAlaAlaThrLeuT 286
                 :::||||  :::::::::
1225 ...................GTCGAAATCCCGGGGCCTTCA 1246

286 hrAlaMetGlyLeuValAspArgAlaGlyArg...............Arg 297
    |||:::      :::||||||||::||||||
1247 TAGCCCTCATCACCATTGACCCGGTGGCCGGCCATCTACCCCATGGCCGTG 1296

298 AlaLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIl 314
    :::  ||||||||||||::||      :::::::      ::|||
1297 TCAAATTGTTGCGGGGGCAGCCTGCCTCGTCATTTATCTCA...... 1341

314 eGlyLeuValSerPheAlaValProMetAspSerGlyProSerCysLeuA 331
    :::                                              }
1341 ..................................................... 1341

331 laValProAsnAlaThrGlyThrGlnThrGlyLeuProGlyAspSerGlyLeu 347
    :::::||                                    :::
1342 ......CCTGAC................................ 1347

348 LeuGlnAspSerSerLeuProProIleProArgThrAsnGluAspGlnAr 364
                                                       }
1347 ............................................... 1347

364 gGluProIleLeuSerGluThrArgAlaLysLysThrLysProHisProArgSerG 381
                                                       }
1347 ............................................... 1347

381 lyAspProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGly 397
                                                       }
1347 ............................................... 1347

398 ProProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLe 414
    |||::|||      :::::::
1348 ......................CTGCACTGGTTAAACATCAT 1367

414 uCysLeuMetValPheValSerAlaPheSerPheGlyPheProValT 431
    ::: |||      :::
1368 AATCATGTGTTTGGGCCGAATGGAATCACCATTGCAATACAAATGATCT 1417

431 hrTrpLeuValLeuSerGluIleTyrProValGlyIleArgGlyArgGlyAla 447
    ||||||  ::|||::::||||||::
1418 GC...CTGGTGAATGCTGAGCTGTACCCCACATTCGTCAGGAAACCTCAGA 1464

448 PheAlaPheCysAsnSerPheAsnTrpAlaAlaAsnLeu.........Ph 461
    :::::|||                                     ||
1465 GTGATGGCTGTGTTCCTCCCTGTGTGACATAGGTGGGATAATCACCCCCCTT 1514

461 eIleSerLeuSerPheLeuAspLeuIleGlyThrIleGlyLeuSerTrpT 478
    |||::: |:::      :::
1515 CATAGTCTTCAGGCTGAGGGAGGTC...................TGGC 1543
```

```
124 rGluLeuValGlyProArgGlnArgGlyValValSerLeuTyrGluA 141
       ::|||||||||::||   ::::::|||
880100 CCGTGAGGCGGCCCGAAGCGGCTGGGCGCCTGGGCGGTGGTGGGCA 880051

141 laGlyIleThrValGlyIleLeuLeuSerTyrAlaLeuAsnTyrAlaLeu 157
       :: :||| ||||:::      ||
880050 TTCCGATGCTGCTCGGCCCGGTCGGTGGCCGATCCTGGTGGTTGGCTC 880001

158 AlaGlyThrProTrpGlyTrpArgHisMetPheGlyTrpAlaThrAlaPr 174
       ||:::: ::||||||||    ::|||   ||
880000 ATCGGGGCC...TACGGCTGCGTTGGATTTC...TTGGTCAACCTGCC 879957

174 oAlaValLeuGlnSerLeuSerLeu......LeuPheLeuProAlaGlyT 189
       |:::: |||::|||   |||   ::|||
879956 GGTCGGGCTGTCCGCGCCTCTTGGCGGGGATCGTGTTCCCAAGAGATC 879907

189 hrAspGluThrAlaThrHisLysAspLeuIleProLeuGlnGlyGlyGlu 205
       |||   ::   |||          |||
879906 GCCCGGCAGCGTCGGAAAACTTCGACTACATGGGCCTCTTG......... 879866

206 AlaProLysLeuGlyProGlyArgProArgTyrSerPhe...LeuAspLe 221
       |||::||||| :::||||||
879865 ......TTGCTGTCGCCCGGCGCTGGCGACCTTCCTGTTCCGGGGTGTCATC 879822

221 uPheArgAlaArgAspAsnMetArgGlyArg............ThrT 233
       ||||||  |||  :::
879821 TAGCCCCGCGTGGAACCGATGGCCGATCGGCACCGTGTTGATACCGGCGA 879772

233 hrValGlyLeuVal.................................. 239
       ::|||||||::||
879771 TCACCGGCCCTGGCGTTGATCGCGGCCATTCGTCGCACATTCGTGTGTACCGC 879722

240 ........................LeuPheGln...GlnLeuThrGl 246
                                ||||||||  :: :::::
879721 ACAGAACATCCGCTCATAGACATGCGCTTGTTCCAGAACCGAGCCGGTCGC 879672

246 yGlnProAsnValLeuCysTyrAlaSerThrIlePheSerSerValGlyP 263
       :||| |||::: ::: |||::::::||
879671 GCAGGCCAACATG...ACGATGACGGTGCTC...TCCCTCGGGC 879634

263 heHisGlyGly.................................... 266
       :: |||::|||
879633 TGTTTGGCTCCTTCTTTGCTGTCTCCCGAGCTACCTCCAGCAAGTGTTGCAC 879584

267 ............SerSerAlaValLeuAlaSerValGlyLeuGlyyAl 278
                   ||||||:::|||
879583 CAATCACCGATGCAATCGGGGGTGCATATCATCCCACAGGGGCCTCGGGTGC 879534

278 aValLysValAlaAlaAlaThrLeuThrAlaMetGlyLeuValAspArgAlaG 295
       :::  ||| :::|||||::::||
879533 C...ATGCTGGCGATCGCCGGAGCGATGATGACCGACGGG 879490

295 lyArgArgAlaLeuLeuAlaGlyLeuCysAlaLeuMetAla......... 308
       :::||| |||::|||
879489 GACCGGCCAAGATGCTGCTGGTTGGGATCATGCTGATCGCTGCGCGGGGTTG 879440

308 ......................................... 308
879...

309 ........LeuSerValSerGlyIleGlyLeuValSerProAlaValAlaValP 322
               :||||:::|||::|||
879389 GCCGGACCGGCTGGCAATCATGGCCATGGCATGGTCTGCTCCATCATGC 879340

322 roMetAspSerGlyProSerCysSerLeuAlaValProAsn......AlaThr 336
       ::||||::|||
879339 CACTGTCCGGCGGCGGCAGTGCAGACCCTGGCCCCACCATCAGATCGCTCGC 879290

337 GlyGlnThr................GlyLeuProGlyAspSerGlyLe 347
       :: |||::|||
879289 GGTTCGACGCTGATCAGCGTCAACCAGCAGGTTGGGCCGTTCGATAGGG... 879242
```

```
347 uLeuGlnAspSerSerLeuProProIleProArgThrAsnGluAspGlnA 364
       :::::||| ::|||
       ACCGCACTGATGTCGGTGCTGCTCACCTACCAGTTCAATC 879202

364 rgGluProIleLeuSerThrAlaLysThrLysProHisProArgSer 380
       ::::: ||:::::||||||||::
879201 ACAGCGGAAATCATCGCTACTGCAAAGAAGTCCACTGACCCCAGAGAGT 879152

381 Gly.............AspProSerAlaProProArgLeuL 390
       |||  |||||||||::::
879151 GGCGCCGGCGGGGGCGGCGGCGGTGACCCTTCCTGCTACCGCCAAAC 879102

390 aLeuSerSerAlaLeuProGlyProGlyProLeuProAlaAArgGlyHisAlaAl 407
       : :::|||
879101 CAACTTCGCGGCC...........................CAAC 879085

407 euLeuArgTrpProThrAlaLeuLeuCysLeuMetValPheValSerArgAlaPhe 423
       |||||::: ::|||||||||||
879084 TGCTGCATGACCTTCGCACGCCTACGCGGTGGTATTCGTGATAGCG... 879038

424 SerPheGlyPheGlyProValThrTrpProValLeuSerGluIleTyrPr 440
       :: :||
879037 ...................ACCCGCGCTAGTGGTCTCGACGCTGATCCC 879009

440 o 440
     |
879008 C 879008
```

```
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-676-967-5

seq_documentation_block:
; Sequence 5, Application US/08676967
; Patent No. 5747317
; GENERAL INFORMATION:
;   APPLICANT: COLLINS, KATHLEEN
;   TITLE OF INVENTION: Human Telomerase
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Science & Technology Law Group
;     STREET: 268 Bush Street, Suite 3200
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94104
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/676,967
;     FILING DATE:
;     CLASSIFICATION: 530
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Osman Ph.D., Richard A
;     REGISTRATION NUMBER: 36,627
;     REFERENCE/DOCKET NUMBER: UCB96-055
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415)343-4341
;     TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2277 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-676-967-5

alignment_scores:
    Quality:  167.50        Length:  687
    Ratio:      0.540       Gaps:     36
```

Fri Mar 15 14:21:22 2002

us-09-652-292-2.rni

Percent Similarity: 45.124   Percent Identity: 25.473

alignment_block:
US-09-652-292-2 x US-08-676-967-5/rev   .

Align seg 1/1   to reverse of: US-08-676-967-5   from: 1 to: 2277

```
  7 ValLeuProLeuCys...AlaSerValSerLeuLeuGlyGlyLeuThrPh  22
    ::|||||||||      :::|||||||||||||||||
2118 CTGCTTCCACTGTTGATCGGGCTTCTTTGGG.....GTGCA 2075

 22 eGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuGlnLeuA  39
    |||||||| ::|||||||||||||||||||||
2074 CGGGCTTCACCTTGCCCTGTCGCG.........CAGGCGG 2043

 39 spPheGlyLeuSerCysLeuGluGln...GluPheLeuValGlySerLeu  54
    ||:::||       |||     :::   |||::::|||      |||
2042 ATCTTGGGGCCGCGGTGGCTGGCAGGGCCAGGCCACCTTGCGGCGCTTCTT 1993

 55 LeuLeuGlyAlaLeu....................................  59
    ::|||    |||
1992 GCCGTCGGGCAGCTCCACCTGCTCCACCTCGGCTTCGTCTTCGGAAGCCGG 1943

 60 ....................LeuAlaSerLeuValGlyPheLeuIleA  70
                       |||||| |||::||||| |||
1942 TCCAGCTGGTGCCGCCGGCCTTGCGCTTCGTCTCGGGGGGCACCTTGCTC 1893

 70 spCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu  86
    ::|||:||||||   |||::::||||||::: |||::||
1892 TGCTCCTCGGT.......GTGGTGCGGCCTTCTGCTGCTGGTC 1852

 87 AlaGlySerLeu.........ThrLeuGlyLeuAla..GlySerLeuAl  99
    ||||||     |||||||||     |||::|||||
1851 CTTGGCGGGCTCGGGCTGGCCCTTCTGGGGCTCCGGCCGGTGGCGGGCTTGC 1802

 99 aTrpLeuValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerS 116
    :::||||   |||::|     |||   ::: |||::::
1801 TGCGCATCTTCTGCGCTGGATGGCAGCTCCTTCCATCTTCAGC 1752

116 erMetAlaCysCys........IleTyrValSerGluLeuValGlyPro 129
    :::|||              |||   |||   |||
1751 TTGCGGCGGTCCTCCAGCTGAACTCCACGATGGGGGCGCTTCAGGGGGCC 1702

130 ArgGlnArgGlyValGlyValLeuValSerLeuTyrGluAla......GlyIleTh 144
    :::::||| ||:::|||    :::::::     |||::|||||:::
1701 GAAGATCTCGGGGTTGTTGTTGATCAGGCCAGGGCCTTCAGGGCGTGCT 1652

144 rVal................GlyIleLeuLeuSerTyrAlaLeuAsnT 155
    |||                |||::: ||||||||:::
1651 CGTGCGTCCTGAACTCGGCGAAGGCGTAGCCCAGGCTCTGGCCCTTCATG 1602

155 yrAlaLeuAlaGlyThrProTrpGlyTrpArgHisMetPheGlyTrpAla 171
    ||::::  :::|||                :::
1601 TTGCCGTGCACGCCCTTCAGGTCGCG......CATCACGCGGCACTCCTT 1558

172 ThrAlaProAlaValLeu...........GlnSerLeuSerLe 182
    |||  |||::::|||             |||:::|||::::
1557 GATGCGGCACGCCCTTCTCGCCGTGGTGCGCCGTGGTGCGCCAGGCTTGCGCCA 1508

182 uLeuPheLeuProAlaGlyThrAspGluThr...................A 193
    |||:::                                             |
1507 GCTGCTTGTCGTCCACGCCCTTGGGCAGGTTGTGCAGGCACAGGCGGGTG 1458

193 laThrHisLysAsp...........LeuLeuProLeuGlnGlyGly 204
    |||||| :::              |||:::  ||||||
1457 CGGCTCACGAAGATGTTCTCGGTCCTTCAGCTTCGTGGTGCTTCAGCAGCTC 1408

205 GluAlaProLysLeuGlyProGlyProArgProTyrSerPheLeuAspLe 221
    ||||||       ||||||       :::::||
1407 GAAGGCGCTCGCGCTTGGCCATGTCGGCGGCCGCTCACGCCCTCGGCCGCCT 1358

221 uPheArgAlaArgAspAsn...MetArgGlyArgThrVal........ 234
    )     ||||::|  :::   :::   |||||||:::  |||
```

```
1357 TGGTGCCGGCGGCGGATCAGGCCCTCGGCGGGCCAGTACAGTTGCGGGTG 1308

235 .....GlyLeu............GlyLeuValLeuPheGlnLeuThr 245
         |||||                ||||||         :::::
1307 CCGGTGGGCGGTTCTTCACCTTGGTGGTCGCAGCTTGGCGGCCTCGTCGCG 1258

246 GlyGlnProAsnVal.............................LeuCysTyrAl 254
    |||:::|||                                   |||  ||
1257 GGTCACGGCCAGGTCCACCTTCAGCTGCGGCGGCCGTCCAGCTTCAGGCCGGC 1208

254 aSerThrIlePheSerSer..............................V 261
    |:::  :::||||||:::
1207 CGGCCTCGTTCTCGGGGCTGGGGCCCAGCAGGACTTCTGGCGGCCTCC 1158

261 alGlyPheHisGlyGlySerSerAlaValLeuAlaSerValGlyLeuGly 277
    ||:||||::::::|||||||||:::::||||||||
1157 TGGGTCATGAACTGGGCAAGGCGCAGCCCTTGCTGTGCTGCGGTGTCGGG 1108

278 AlaValLysValAlaAlaThrLeuThr...........................A 287
    :::        |||||:::|||
1107 GTGCAGCACGATGCGCACGTACTTCAGCTCGCCGAACTGCTCGCAGCAGCT 1058

287 ...................................................A 287
                                                       .
1057 CGCCCAGCTCCTCCTCCTCGCTGTGAAGCTCAGTTGCCGATGAACACG 1008

287 laMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLeu....... 301
    :::::::||||||    ||||||:::  |||||||||||
1007 GTCTTGCCCTCGTTCACGTCGCTGGGCAGCTTGGCGTTCTTCTTGTTGCT 958

302 .............................AlaGlyCysAlaLeu. 306
                                 ||:|||  :||||
957 CACCTTGCACGGCCTTGTCTCCTCGCTCCTCGGTCGCTGGCTGTCTGGG 908

307 ....MetAlaLeuSerValSerGlyIleGlyIleLeuValSerPheAlaValP 322
        :::                 ||||||         |||||
907 CCAGCTCCTCCGCCGTCGTGATGCTGTCGCCGTTGGCGTTCTTCTTGTGTC 859

322 roMetAspSerGlyProSerCysLeuAlaValProAsnAlaThrGlyGln 338
    |||||||           :::CTCCTCGTGTGGTCGCTGCTCTTGGCGGGG 829
858 CTCCTCGTGTGGTCGCTGCTCTTGGCGGGG 829

339 ThrGly..........................LeuProGlyAspSerGlyLe 347
    |||                            ||| ||| :::   |||
828 GGGCGGCCGCTTCACGGCGCGCTTCTGGATCTGGACGGCCTTGGTCACCT 779

347 uLeuGlnAspSerLeuProProIleProArgThrAsnGluAspGlnA 364
    |||   ||||||   |||||| ||||||       :::::
778 TGCTCTCGATGT.TCTCCTCCTCCTCGT...CGTCGAACACGCC 733

364 rgGluProIleLeuSerThrAlaLysThrLysProHisProArgSer 380
    |||                :::|||  |||       |||||||
732 GTCCTCCTCGTCGTCGTCGTCGTCGTCGTCGTCGTTCTCCTCCT. 689

381 GlyAspProSerAlaProProArgLeuAlaLeuSerAlaLeuProGl 397
    ||||||                |||||||||
688 ...CCTCCATGTCCTCCTCCTCGCGGCCCTTCTTCTTCACGCTCTCCTGG 642

397 yProProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAla.Leu 413
    :::|||||||             |||||| |||:::
641 T......GCTTGCTCGTGGCTCTTCCTCCGCCGATGGCGGCTCACGCT 598

414 LeuCysLeuMetValPheValSerAlaPheSerPheGlyPheGlyProVa 430
    |||   |||:::|||:::|||        :::||| |||||||||
597 CTGGGGTCCCTTGTACTTGTC.......CTTGGGCCACGGCGGCCCAGT 560

430 l.ThrTrpLeuVal.................................... 434
    .|||||||||
559 CCACGGCCACGGTGCGGCGCCCTTGATCTCCTTCATGTTCATGCCCTTCAGG 510

435 ...........................................LeuSerGl 437
                                               :::
509 GCCTTGCCGGCCTCCAGCAGGTTCTTGAACTGACGAAGCCGAAGCCGCG 460
```

```
                                Ratio:  0.540                          Gaps:        36
Percent Similarity:   45.124          Percent Identity: 25.473

alignment_block:
US-09-652-292-2 x US-08-676-974-5/rev

Align seg 1/1   to reverse of: US-08-676-974-5   from: 1   to: 2277

    7 ValLeuProLeuCys...AlaSerValSerLeuLeuGlyGlyLeuThrPh 22
      ::|||||||||          :::|||||||||          |||
 2118 CTGCTTCCACTGGTGATCTGGGCTTGGGCTTCTTGGG......GTGCA 2075

   22 eGlyTyrGluLeuAlaValAlaIleSerGlyAlaLeuProLeuGlnLeuA 39
      |||  :::||||||||::::::::          :|||  |||
 2074 CGGGCTTGGGCTCCTTCTTGGGGCACTCG.....CTGTTCTCGTTCTTGCCCT 2043

   39 spPheGlyLeuSerCysLeuGluGln...GluPheLeuValGlySerLeu 54
      ||:::|||    |||||| |||  ::   |||::::::   |||
 2042 ATCTTGGGCGCCGGTGGCTGGGCAGGGCCAGCACCTTGCGGCGCCTTCTT 1993

   55 LeuLeuGlyAlaLeu.......                              59
      ::::|||  |||
 1992 GCCGTCGGCAGCTCCACCTGCTCCACCTCGGCCTTGGTCTCGGAAGCCGG 1943

   60 .......................LeuAlaSerLeuValGlyGlyPheLeuIleA 70
                           ||||||  |||::||||::|||||:  |||
 1942 TCCAGCTGGTGCGCGGCCTTGCGCTTCTGCTCGGGGGGCACCTTGCTC 1893

   70 spCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu 86
      |||      :::::|||||||||:  :|||::::|||:::|||:|:|||::
 1892 TGCTCCTCGGT..........GTGGTGCTGGGCGGCCTTCTGCTGCTGGTC 1852

   87 AlaGlySerLeu.........ThrLeuGlyLeuAla..GlySerLeuAl 99
      |||||||||||          ||||::::|||||||  ::|||:|||||
 1851 CTTGGCGGCGTCGGGCTCGGGCCTGGCCCTTCTGGGGGCTCGCCGGGCTTGC 1802

   99 aTrpLeuValLeuGlyArgLeuValArgAlaValGlyPheAlaIleSerLeuSers 116
      :  :||||  |||                      |||::  :|||||||  :::
 1801 TGCGCATCTTCTGCAGGCTGCCTGGATGCGCAGCTCCTTCATCTTCAGC 1752

  116 erMetAlaCysCys.......              IleTyrValSerGluLeuValGlyPro 129
      |||                               ::  |||   ||| |||:::
 1751 TTGCGGCGGTCCTCCAGGCTGAACTCCACGATGGGGCCTTCAGGGGGCC 1702

  130 ArgGlnArgGlyValLeuValSerLeuTyrGluAla......GlyIleTh 144
      |||::: |||||||:          ||||||::::|||    |||:::
 1701 GAAGATCTCGGGGTTGTTGTTGATCAGGCGACGGGCCTTCAGGGCGTGCT 1652

  144 rVal........................              GlyIleLeuLeuSerTyrAlaLeuAsnT 155
      |||                                       ||||||:::|||
 1651 CGTGCTCCTGGAACTCGGCGAAGGCGTAGCCCAGGCTCTGGCCCTTCATG 1602

  155 yrAlaLeuAlaGlyThrProTrpGlyTrpArgHisMetPheGlyTrpAla 171
      ||:::: :::                   |||        |||||||:::
 1601 TTGCCGTGCACGCCCTTCAGGTCGCG......CATCACGCGGCACTCCTT 1558

  172 ThrAlaProAlaValLeu.........              GlnSerLeuSerLe 182
      ||:::::|||:::|||                         |||:::|||::|||
 1557 GATGCGCACGCCCCTTCTCCGCCGCTGGTGCGGCCTCAGCAGCTTGCGCA 1508

  182 uLeuPheLeuProAlaGlyThrAspGluThr.................A 193
      ||:::::|||                                       |
 1507 GCTGCTTGTCGTCGCACGGCCTTGGGCAGGTTGTGCAGGCACAGGCGGGTG 1458

  193 laThrHisLysAsp.............              LeuIleProLeuGlnGlyGly 204
      ||:::|||::|||                            |||::  |||
 1457 CGGCTCACGAAGATGTTCTGGTCCTTCAGCTTCTGGTGCTTCAGCAGCTC 1408

  205 GluAlaProLysLeuGlyProGlyArgProArgTyrSerPheLeuAspLe 221
      |||||||                      :::||||::||||||
 1407 GAAGCGGCTCGCGCTTGGCCATGTCGGCCGGCGCTTCACGCCCTCGGCGCCT 1358

  221 uPheArgAlaArgAspAsn...MetArgGlyArgThrThrVal........ 234
```

```
509 GCCTTGCCGCGCCTCCAGCAGGTTCTTGAACTGACGAGAAGCCGAAGCCGCG 460
437 uIeTyrProValGluIleArgGlyArgAlaPheAlaPheCysAsnSerP 454
         :||:|||    ||| ::|||:|    ::||:::
459 CATCTTGCCGTCGGGCTTGGCGGGG...ATGTTCACCTCCAGCACGGCGC 413
454 heAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPhe........... 466
         |||||||| ::::::|||   |||:::
412 CGAACTGGGCGAACACGGTCTTCAGGTCGTCCTGCCACTGAAGCTC 363
467 ...LeuAspLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTy 482
         |||  :::|||    ||||| :::
362 AGGTTGCGGATGATCAGGCGGGCCTTCTTGTCGGCCACCTTGGCCTTCTT 313
482 rGlyLeuThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProG 499
         ::: |||||    |||||    |||::|||:::||
312 GGGCTTGGGCTCCTTCTTGGGCACTCG...CTGTTCTCGTTCTTGCCCT 266
499 luThr..........LysGlyGlnSerLeuAlaGluIleAspGlnGln 511
         ||||||        :::::
265 TCTCCTTGGTCTTGTTGCGCAGCTTCTTCTTGGCCACGGTCACGTTGATC 216
512 PheGlnLysArgArgPheThrLeuSerPheGlyHisArgGlnAsnSerTh 528
         :::|||   ::: :::::: ||||| |||   ::::|| ::|||
215 TTGCAGCCCTCGAAGGTGGTCATCTCCTTCAGGGCGCCGCTGCACGTCCTC 166
528 rGlyIle 530
         ::::::
165 CAGCATG 159
```

alignment_scores:

```
1357 TGGTGCCGGCGCGGATCAGGCCCTCGCGGCCAGGTACAGGTTGCGGGGTG 1308
          :::    |||||   :::    |||   |||
 235 ...GlyLeu..........GlyLeuValLeuPheGlnLeuThr 245
          |||||       |||       ::::::
1307 CCGGTGGGCTTCTTCACCTTGGTCGTCGGCCGGCCTCCTCGCG 1258
 246 GlyGlnProAsnVal............LeuCysTyrAl 254
          |||::: ::|||     |||  ||
1257 GGTCACGGCCAGGTCCACCTTCAGCTTCAGGCCGC 1208
 254 aSerThrIlePheSer..............    V 261
          |::|| ::|||          :
1207 CGGCCTCGTTCTCGGGGCTGGGCGGCCAGCAGGCACTTCTGGGCGGCCTCC 1158
 261 aIGlyPheHisGlySerSerAlaValLeuAlaSerValGlyLeuGly 277
          :::||| :::||||| :::|||:::|||
1157 TGGGTCATGAACTGGGCAAGGCGCAGCCCCTTGCTGTCGTGCTGGTGTCGGG 1108
 278 AlaValLysValAlaAlaThrLeuThr.......... 286
          :::    |||    ::|||
1107 GTGCAGCACGATGCCGACGTACTTCAGCTCGCCGAACTGCTGCAGCAGCT 1058
 287 .............................A 287
                                      :
1057 CGCCCAGCTCCTCCTCCGCGTCGAAGCTCAGGTTGCGGATGAACACG 1008
 287 laMetGlyLeuValAspArgArgAlaGlyArgArgAlaLeuLeuLeu...... 301
          ::::::    |||||||||
1007 GTCTTGCCCTTCGTTCACGCCGCTGGGCAGCTTGCCGCTTGCCGTCACCT 958
 302 .................AlaGlyCysAlaLeu.. 306
                           ||||| |||||||
 957 CACCTGCACGGCCTTGTCCTCCGCGCCTCCTGCGGTGTCGCTCTCGG 908
 307 ...MetAlaLeuSerValSerGlyIleGlyLeuValSerPheAlaValP 322
          :::   |||:::::::|||   ::|||::::  |||||
 907 CCAGCTCCTCGCGCGTCGATGCTGTCCTCCTCCAGGTCGCTGTC. 859
 322 roMetAspSerGlyProSerCysLeuAlaValProAsnAlaThrGlyGln 338
          |||||||||   ::   |||
 858    CTCCTCGGTCGTGGTCGCTGCTCTTGGCGGG 829
 339 ThrGly..............LeuProGlyAspSerGlyLe 347
          |||       :::    |||||||
 828 GGCGGCGGCCTTCACGGCCGCTTCGGATCTGCACGGGCTTGGTCACCT 779
 347 uLeuGlnAspSerLeuProProIleProArgThrAsnGluAspGlnA 364
          |||   :::::|||    ||||||
 778 TGCTCTCGATGT.TCTCCTCCTCCGCTCCTCGT...CGTCGAACACGCC 733
 364 rgGluProIleLeuSerThrAlaLysThrLysProHisProArgSer 380
          ||||||||    |||       |||
 732 GTCCTCCTCGCGCGTCGTCGCGTCCGTCGTCCGTTCTCCTCCT...... 689
 381 GlyAspProSerAlaProArgLeuAlaLeuSerSerAlaLeuProGl 397
          ||||||| ||||||
 688 ..CCTCCATGTCCTCTCCCTCCGGCGCCCTTCTTCTTCACGCCTCCTGG 642
 397 yProProLeuProAlaArgGlyHisAlaLeuLeuLeuArgTrpThrAla..Leu 413
          :::||| |||||||
 641 T......GCTTGCTCGTGGCTCTCTCCCGCCGATGGCGTCACGCT 598
 414 LeuCysLeuMetValPheValSerAlaPheSerGlyPheGlyProVa 430
          |||:::::||:||||   ||||||||
 597 CTGGGGTGTCCTTGTACTTGTC...........CTTGGCCACGGCGCCCAGT 560
 430 lThrTrpLeuVal.......... 434
          |||   |||
 559 CCACAGGCCACGGTGCGGCCCCTTGATCTCCTTCATGTTCATGCCCTTCAGG 510
 435 ..................LeuSerGl 437
                            ::
```

```
          Quality:      167.50         Length:    687
          Ratio:        0.540          Gaps:       36
Percent Similarity:     45.124         Percent Identity: 25.473

alignment_block:
US-09-652-292-2 x US-09-098-487-5/rev

Align seg 1/1  to reverse of: US-09-098-487-5   from: 1  to: 2277


   7 ValLeuProLeuCys..AlaSerValSerLeuLeuGlyLeuGlyLeuThrPh  22
     ::::|||||||||   ::::::|||||||||
2118 CTGCTTCACTGGTTGATCTGGGCTTGGGCTTCTTGGG.....GTGCA 2075

  22 eGlyTyrGluAlaValIleSerGlyAlaLeuLeuProLeuGlnLeuA  39
     |||  ::|||||||::::::
2074 CGGGCCTTCACCTTGCCCTGTCGCG.............CAGGCGG 2043

  39 spPheGlyLeuSerCysLeuGluGln...GluPheLeuValGlySerLeu  54
     ||:::|||  |||   :::  :||::::  |||
2042 ATCTTGGGCCGGTGCTGGGCAGGGCCAGCACCTTGCGGCGCTTCTT 1993

  55 LeuLeuGlyAlaLeu...........................  59
     ::::|||
1992 GCCGTCGGGCAGCTCCACCTGCTCCACCTCGGCCTTGGTCTGGAAGCCGG 1943

  60 ..........LeuAlaSerLeuValGlyLyGlyPheLeuIleA  70
               ||||||   |||:::||||||   |||
1942 TCCAGCTGGTCGCCGCGGTTCTGCGCTCGGGGGGCCACCTTGCTC 1893

  70 spCysTyrGlyTyrArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeu  86
     :::|||       :::||||||:::  ::::::||:::
1892 TGCTCCTCGGT.......GTGGTGCGTGGGGCGCCCTTCTCTGCTGGTC 1852

  87 AlaGlySerLeu.......ThrLeuGlyLeuAla...GlySerLeuAl  99
     |||::::|||         ||||||||||   |||:::||||
1851 CTTGCGGGCTCGGGCTTGGGCCTTGGGGCTGCCCGGTGGGGCGGCCTTGC 1802

  99 aTrpLeuValLeuGlyAlaArgAlaValAlaValGlyPheAlaIleSerLeuSers 116
     :::|||       :::|||   :::      ::||::
1801 TGCCATCTTCTGCGAGGCTGCGCTGGATGGCGAGCTCCTTCATCTTCAGC 1752

 116 erMetAlaCysCys.......IleTyrValSerGluLeuValGlyPro 129
     |||:::   :::||||||       ||| |||
1751 TTGCGGCGTCCTCCAGGCTGAACTTCCACGATGGGGGCCTTCAGGGGCC 1702

 130 ArgGlnArgGlyValLeuValSerLeuTyrGluAla......GlyIleTh 144
     :::::::|||||  :::|||:::::                  :||:::
1701 GAAGATCTCGGGGTTGTTGTTGATCAGGCGCAGGGCCTTCAGGGCGTGCT 1652

 144 rVal..............GlyIleLeuLeuSerTyrAlaLeuAsnT 155
     |||               |||:::   ::: ||||||:::
1651 CGTGTCCTCGAACTCGGCGAAGGCGTAGCCCAGGCTCTGGCCCTTCATG 1602

 155 yrAlaLeuAlaGlyThrProTyrTrpGlyTrpArgHisMetPheGlyTrpAla 171
     :::      |||  :::||| :::  |||
1601 TTGCCGGTGCACGCCCTTCAGGTGCGG.....CATCACGGCGCACTCCTT 1558

 172 ThrAlaProAlaValLeu..................GlnSerLeuSerLe 182
     |||  |||:::|||||                    :::|||:::
1557 GATGCGGACGCGCCCTTCTCGCCGGCCTGGTGCGCCTGCAGCAGCTTGCGCA 1508

 182 uLeuPheLeuProAlaGlyThrAspGluThr.............A 193
     |||:::  |||      :::
1507 GCTCGTTGTCGTCCACGCCTTGGGCAGGTTGTGCAGGCACAGGCGGGTG 1458

 193 laThrHisLysAsp............LeuIleProLeuGlnGlyGly 204
     :::|||:::|||              ||||| |||
1457 CGGCTCACGAAGATGTTTCTGCCTTCAGCTTCTGGTGCTTCAGCAGCTC 1408

 205 GluAlaProLysProGlyProGlyArgProArgTyrSerPheLeuAspAsp 221
     :::::::|||    :::||||||:::::::::
1407 GAAGCGGCTCGCCGTTGGCCATGTCGGCGGCGCGCCTCCGCGCCCT 1358


 221 uPheArgAlaArgAspAsn...MetArgGlyArgThrThrVal.......  234
     :::   |||::   |||   ::: |||                         111
1357 TGGTTCCGGCGCGGATCAGGCCCTCGCGGCCAGGTACAGGTTGCGGGTG 1308

 235 ....GlyLeu..........GlyLeuValLeuPheGlnGlnLeuThr  245
     ||||||          |||
1307 CCGGTGGGCTTCTTCACCTTGGTGGTCTGCAGCTTGGGCGCCTCGCCG 1258

 246 GlyGlnProAsnVal..................LeuCysTyrAl  254
     ||:::   ::|||                    |||
1257 GGTCACGGCCAGGTCCACCTTCAGCTTGGGCCGTCCAGCTTCAGGCGC 1208

 254 aSerThrIlePheSerSer..................V  261
     |:::  :::::::::|||
1207 CGGCCTCGTTCTCGGCGGTGGCGGCGGCCAGCAGCACTTCTGGGCGGCCTCC 1158

 261 aIGlyPheHisGlyIleGlySerSerAlaValLeuAlaSerValGlyLeuGly  277
     ::|||   |||:::::|||||| :::||::|||
1157 TGGGTCATGAACTGGCGAAGGCCAGCCCTTGCTTGCTGCTCGGTGTCGGG 1108

 278 AlaValLysValAlaAlaThrLeuThr...........  286
     :::    :::|||
1107 GTGCAGCACGATGCCACGTACTTCAGCTCGCCGAACTGCTGCAGCAGCT 1058

 287 ...........................................A  287
1057 CGCCCAGCTCCTCCTCGCTGCGAAGCTCAGTTGCGGATGAACACG 1008

 287 laMetGlyLeuValAspArgAlaGlyArgAlaArgAlaLeuLeu......  301
     :::::::|||||  |||||||||||||||||||
1007 GTCTTGCCCTGCTTCACGTCGTGGGCAGCTTGGCTTCTTCTTGTTGCT 958

 302 ...............................AlaGlyCysAlaLeu.  306
                                    ||||||:::||:||||
 957 CACCTGCACGGCCTTGTCTCCTCCTGCTCCTCGGTGTCGGTGTCGCTCTGGG 908

 307 ....MetAlaLeuSerValSerGlyIleGlyLeuValSerPheAlaValP 322
     :::     |||:::|||||::::::|||:::
 907 CCAGCTCCTCCGCGTCGCGTCGATGCTGTCGCCTCTCCTCCAGGTCGCTGTC. 859

 322 roMetAspSerGlyProSerCysLeuAlaValProAsnAlaThrGlyGln 338
     ||||||||||                CTCCTGCTGTGGTCGCTCTTGGCGGG
 858 GTCCTCCCTCGTCGTGGTCGTCGTCGTGGTTCTCCTCCT         829

 339 ThrGly........................LeuProGlyAspSerGlyLe 347
                                   ||| :::::|||   ||||
 828 GGCGGGGGCGCTTCACGGCGGCGCCTTCGGGATCTGCACGGGCGGTTGTCACCT 779

 347 uLeuGlnAspSerSerLeuProProIleProArgThrAsnGluAspGlnA 364
     |||   |||||||  :::   ||||| :::
 778 TGCTCTCGATGT..TCTCCTCCTCCTCGTCCTCGT...CGTCGAACACGCC 733

 364 rgGluProIleLeuSerThrAlaLysThrHisLysProHisProArgSer 380
     ::|||||||||||||::::::|||
 732 GTCCTCCCTCGTCGTCACGCCTCGCCGTCGTCGTCGTTCTCCTCCT 689

 381 GlyAspProSerAlaProProArgLeuAlaLeuSerSerAlaLeuProGl 397
     :::::||| ||||||||||||||||||||||
 688 ...CCTCCATGTCCTCCTCTCGCGGCCGCCCTTCTCTTCCACGCTCTCTGG 642

 397 yProProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAla.Leu 413
     |||   |||||||::|||       ::::::||||::::||::  :::
 641 T.......GCTTGCTCTCGTGGCCTCTTCTCCTCGCCGATGGCGCTCACGCGT 598

 414 LeuCysLeuMetValPheValSerAlaPheSerAlaPheGlyProAla 430
     |||     |||:::|::||:::::
 597 CTGGGTGTCCTTGTACTTGTC............CTTGGCCACGGCCCCAGT 560

 430 l.ThrTrpLeuVal......................  434
     :|||  |||
 559 CCACGGCCACGGTGCGGCCCCATGTCGGCGGGCCGCGCCTTGATCTCCTTCATGTCCTTTCATGTCCCTTCAGG 510

 435                    LeuSerGl 437
```

```
509 GCCTTGCCGCCTCCAGCAGGTTCTTGAACTGCACGAAGCCGAAGCCGCG 460
                                                  ::

437 uIleTyrProValGluIleArgGlyArgAlaPheAlaPheCysAsnSerP 454
    :|||  |||      ::|||||||      ||||:::      :::::::
459 CATCTTGCCGTCGGGCTTGCGGGGG...ATGTTCACCTCCAGCACGGCGC 413

454 heAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPhe............ 466
    ||||||||| :::::|||  |||  ||| |||::
412 CGAACTGGGCGAACACGGTCTTCAGGTCGTCCTCGCTGCACTTGAAGCTC 363

467 ...LeuAspLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTy 482
       |||  :::|||      ||||||  |||||  :::
362 AGGTTGCGGATGATCAGGCGGGCCTTCTTGTCGGCCACCTTGGCCTTCTT 313

482 rGlyLeuThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProG 499
    :::||| :::  )||||| ::: :::::::||  |||:::|||
312 GGCCTTGGGCTCCTTCTTGGGGCACTCG...CTGTTCTCGTTCTTGCCCT 266

499 luThr..........LysGlyGlnSerLeuAlaGluIleAspGlnGln 511
    :::|           :::::  ||||||     :::
265 TCTCCTTGGTCTTGTTGCGCAGCTTCTTGGCCACGGTCACGTTGATC 216

512 PheGlnLysArgArgPheThrLeuSerPheGlyHisArgGlnAsnSerTh 528
    :::  |||      :::::::|||||||  |||  ::|||::
215 TTGCAGCCCTCGAAGGTGGTGATCTCCTTCAGGGCGGCGCACGTCCTC 166

528 rGlyIle 530
    :::::::
165 CAGCATG 159
```